МІЖРЕГІОНАЛЬНА
АКАДЕМІЯ
УПРАВЛІННЯ ПЕРСОНАЛОМ

Р. М. Трохимчук

ДИСКРЕТНА МАТЕМАТИКА

МІЖРЕГІОНАЛЬНА
АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ

МАУП

Р. М. Трохимчук

ДИСКРЕТНА МАТЕМАТИКА

*Рекомендовано
Міністерством освіти і науки України
як навчальний посібник
для студентів вищих навчальних закладів*

Київ
ДП «Видавничий дім «Персонал»
2010

ББК 22.176 я73
Т76

Рецензенти: *В. Н. Редько*, д-р фіз.-мат. наук, проф., акад. НАН України
С. Л. Кривий, д-р фіз.-мат. наук, проф.

Схвалено Вченою радою Міжрегіональної Академії управління персоналом (протокол № 5 від 25.05.04)

Рекомендовано Міністерством освіти і науки України (лист № 14/18.2-1024 від 20.05.04)

Трохимчук Р. М.

Т76 Дискретна математика: Навч. посіб. для студ. вищ. навч. закл. — К.: ДП «Видавничий дім «Персонал», 2010. — 528 с.: іл. — Бібліогр.: с. 504–506.

ISBN 978-966-608-863-8

У навчальному посібнику викладено основи дискретної (комп'ютерної) математики, складової теоретичної (математичної) кібернетики — науки, що займається розробкою та дослідженням комп'ютерних і програмних систем. Матеріал проілюстровано численними прикладами. До кожного розділу подано задачі та вправи, що дає змогу використовувати цей посібник для практичних занять і самостійної роботи. До більшості задач наведено розв'язки чи вказівки.

Для студентів вищих навчальних закладів і всіх, хто бажає здобути фундаментальні знання в галузі сучасної інформатики.

ББК 22.176 я73

ISBN 978-966-608-863-8

ПЕРЕДМОВА

Найвидатніший винахід XX століття — це, безумовно, електронна обчислювальна машина, або комп'ютер. Широка екстенсивна й інтенсивна комп'ютеризація всіх сфер людського буття — виробництва, економіки, наукових досліджень, освіти, побуту — основний напрям інтенсифікації фізичної й інтелектуальної праці людини, підвищення ефективності та продуктивності праці, запорука прискорення науково-технічного й соціального прогресу.

Математика та математики відіграли видатну й визначальну роль як у формулюванні та розвитку абстрактної ідеї комп'ютера (А. Тьюрінг, Е. Пост, А. Черч, С.-К. Кліні, А. А. Марков та ін.), так і в процесі практичної реалізації обчислювальної машини (Дж. фон Нейман, С. О. Лебедєв, В. М. Глушков та ін.).

Без перебільшення математику можна вважати "законним батьком" комп'ютера, батьком уважним, чуйним і відповідальним. І зразу після народження, і тепер, коли "дитина" подорослішала й вирушила в самостійну подорож світами, математика не залишає її поза своєю пильною увагою. Вона вчила свій винахід робити перші кроки (монополізувавши, до речі, використання та "забезпечення роботою" перших комп'ютерів), повноцінно спілкуватися з навколишнім середовищем і подібними собі об'єктами, забезпечила йому добру освіту, постійно дбає про підвищення його кваліфікації.

У свою чергу, комп'ютер не міг не вплинути на саму математику. З його появою відбулася певна передислокація розділів класичної математики та перерозподіл уваги й зацікавленості математиків до предмета досліджень. Зокрема, з'явився окремий розділ сучасної математики, що дістав назву *дискретної математики*.

Для жодної наукової дисципліни не існує (і в принципі не може існувати) чіткого та повного означення її предмета й меж відповідних наукових досліджень. З одного боку, це можна пояснити постійним і непередбачуваним розширенням сфер досліджень, неможливістю точно визначити всі проблеми, напрями, інтереси та зміст будь-якої галузі науки. З іншого боку, ніколи не припинявся, а останнім часом інтенсифікувався процес взаємовпливу та взаємопроникнення різних, спочатку досить далеких одна від одної, наукових теорій і дисциплін.

У довідниках, енциклопедіях і вступних розділах підручників і монографій означення предмета певної наукової дисципліни (фізики, хімії, біології тощо) має описовий характер; зазвичай його подають за такою схемою. Спочатку формулюють у найзагальніших рисах основні напрями досліджень і методологію галузі науки. Як приклад можна навести таку цитату з енциклопедії: "Фізика — наука, яка вивчає найпростіші й водночас найзагальніші закономірності явищ природи, властивості й будову матерії та закони її руху... Фізика належить до точних наук і вивчає кількісні закономірності явищ". Далі подають перелік основних розділів, об'єднаних у єдину наукову дисципліну згідно з традицією чи за спільною згодою науковців (авторитетів у цій науці).

Якщо скористатися зазначеною схемою, то першу частину означення можна сформулювати так: *дискретна математика* (інші назви: *дискретний аналіз*, *скінченна математика*) — це розділ сучасної математики, у якому вивчають властивості математичних об'єктів дискретного характеру, тобто таких, основні параметри яких мають скінченні чи зліченні області означення й області зміни (значень).

Хоча окремі розділи дискретної математики виникли в далекій давнині, вона остаточно сформувалась як окрема математична дисципліна лише у XX столітті. Тому ще не створено вищезгаданої традиції, остаточно не досягнуто спільної згоди фахівців щодо багатьох розділів дискретної математики, і в пропонованих різними авторами переліках розділів, що складають її предмет, є істотні відмінності. У цьому можна переконатися, порівнявши зміст деяких підручників або навчальних посібників з дискретної математики.

Водночас більшість авторів погоджуються, що основним стимулом до об'єднання різних розділів класичної математики — тих, що вже існували здавна, і нових — під терміном "дискретна математика" стало виникнення та широке впровадження електронних обчислювальних машин, а також поява однієї з найважливіших наук XX століття — кібернетики. Дискретна математика стала складовою частиною кібернетики, точніше ядром її теоретичної (математичної) частини — теоретичної (математичної) кібернетики, науки, що активно й потужно розвивається.

Нині бракує літератури, яку могли б використовувати викладачі та студенти для підготовки навчальних курсів і вивчення дискретної математики. За останні три десятиліття вийшло декілька книг, у назві яких є словосполучення "дискретна математика" чи які мають відношення до цієї дисципліни. Однак велика кількість цих видань за своїм змістом орієнтована на своєрідне та дуже специфічне уявлення про предмет дискретної математики, яке умовно можна назвати "московським", оскільки його головними ідеологами й авторитарними провідниками були московські математики О. Б. Лупанов і С. В. Яблонський, а також їхні послідовники й учні — В. А. Горбатов, В. Б. Кудрявцев, Л. А. Шоломов [10, 38] та ін.

Істотно інше поняття про предмет і зміст дискретної математики склалося в Україні, де безперечним авторитетом у його визначенні були академік В. М. Глушков і його колеги й учні (Л. А. Калужнін, К. Л. Ющенко, В. Н. Редько, О. А. Летичевський та ін.). На жаль, це поняття не знайшло відображення в одній окремій книжці, а розпорошене в декількох цікавих і змістовних виданнях, які на сьогодні малодоступні (див., наприклад, [7, 9] та ін.).

Головним критерієм, за яким різні розділи математики (як ті, що виникли в давні часи, так і ті, що з'явились у XX столітті) об'єднують під назвою "дискретна математика", є, безумовно, їхнє відношення до теорії та практики проектування й використання електронних обчислювальних машин і програмування. Тому ще однією назвою дискретної математики є "*комп'ютерна математика*", що засвідчила й видана в 1984 р. книга американських авторів Д. Кука та Г. Бейза з такою назвою [18], і зміст якої цілком узгоджується із загальноприйнятим розумінням предмета дискретної математики. На наш погляд, останній термін є найбільш відповідним для даної дисципліни.

Основна мета цього підручника — викладення фундаментальних базових понять, принципів, результатів і методів із царини комп'ютерної математики, підготовка для подальшого ширшого та глибшого вивчення предмета.

Темпи розвитку комп'ютерного господарства (як його матеріальної "залізної" частини, так і особливо програмного забезпечення) такі, що знання, отримані під час навчання, часто застарівають вже на момент його завершення. Тут мова йде про знання та навички володіння конкретними системами й розробками. Не девальвуються лише фундаментальні базові знання. Тому фахівець, який бажає завжди залишатися якщо не "на висоті", то принаймні "на рівні" ідей, новацій і прогресу, має подбати про фундаментальну та глибоку теоретичну підготовку.

Основу цього підручника склали лекції з нормативного курсу "Дискретна математика", протягом 1991–2004 рр. читаного студентам факультету кібернетики Київського національного університету імені Тараса Шевченка.

Майже до всіх розділів підручника підібрано задачі та вправи, кількість і рівень яких достатні для того, щоб використовувати цей підручник також для проведення практичних занять. Для більшості задач наведено розв'язки чи вказівки, що істотно полегшить самостійне опрацювання матеріалу. Додаткові задачі можна знайти в збірниках [4–6, 20, 27–30, 34].

Автор із вдячністю й увагою готовий вислухати відгуки, зауваження, побажання, поради, інформацію про помічені помилки чи недоречності, які можна надсилати на адресу: 03022, Київ-22, пр. Акад. Глушкова, 6, факультет кібернетики, або електронною поштою: trost@unicyb.kiev.ua.

РОЗДІЛ 1

МНОЖИНИ ТА ВІДНОШЕННЯ

Основи теорії множин було закладено відомим німецьким математиком Георгом Кантором у другій половині дев'ятнадцятого століття. Появу цієї теорії з ентузіазмом сприйняли багато авторитетних математиків. Вони побачили в ній засоби для створення метамови математики, тобто формальної універсальної системи понять і принципів, за допомогою якої можна було б викласти з єдиних позицій зміст різноманітних традиційно далеких один від одного розділів математики. Перші такі досить успішні спроби було зроблено вже незабаром після виникнення канторівської теорії множин.

Однак пізніші дослідники виявили в теорії Кантора чимало суперечностей — так званих *парадоксів*, або *антиномій*, теорії множин. Виникла кризова ситуація. Одна частина математиків, посилаючись на штучність сформульованих антиномій, вважала за краще не помічати цих суперечностей або не надавати їм великого значення. Водночас інша (скажімо, відповідальніша) група математиків зосередила свої зусилля на пошуках більш обґрунтованих і точних принципів і концепцій, на яких можна було б побудувати несуперечливу теорію множин.

У результаті було запропоновано кілька формальних (аксіоматичних) систем, які служать фундаментом сучасної теорії множин, а значить, і всієї класичної математики. Важливість цих досліджень серед іншого підкреслює й той факт, що значний внесок у становлення аксіоматичної теорії множин зробили такі видатні математики й мислителі минулого століття, як Б. Рассел, Д. Гільберт, К. Гедель та ін.

Сьогодні теорія множин — це математична теорія, на якій ґрунтуться більшість розділів сучасної математики, як неперервної, так і дискретної.

Докладніше з історією виникнення та розвитку теорії множин можна ознайомитися, прочитавши цікаву монографію А. Френкеля та І. Бар-Хіллела "Основи теорії множин" [32] чи книгу М. Клайна "Математика. Втрата певності".

1.1. Поняття множини. Способи задання множин

Для наших цілей достатньо буде викласти основи так званої *інтуїтивної*, або *наївної*, теорії множин, яка в головних своїх положеннях зберігає ідеї та результати засновника теорії Г. Кантора.

В інтуїтивній теорії множин поняття "***множина***" первинне й неозначуване, тобто його не можна означити через інші простіші терміни чи об'єкти, а можна тільки пояснити на прикладах, апелюючи до нашої уяви й інтуїції. У математиці є й інші такі поняття: "число", "пряма", "точка", "площина" тощо.

Канторівський вираз: "Множина — це зібрання в єдине ціле певних об'єктів, чітко розрізнюваних нашою інтуїцією чи нашою думкою", — безумовно, не можна вважати строгим математичним означенням. Це, скоріше, пояснення поняття множини, яке заміняє термін "множина" на термін "зібрання". Інші синоніми основного слова "множина" — "сукупність", "набір", "колекція", "об'єднання" тощо.

Прикладами множин можуть бути множини десяткових цифр, літер українського алфавіту, мешканців Києва, парних чисел, розв'язків якогось рівняння та ін.

На письмі множини позначають зазвичай великими літерами. Для деяких множин у математиці використовують сталі позначення, наприклад: Z — множина цілих чисел, N — множина натуральних чисел (зокрема, через N_k будемо позначати множину натуральних чисел від 1 до k), Q — множина раціональних чисел, R — множина дійсних чисел тощо.

Об'єкти, з яких складається задана множина, називаються її ***елементами***. Елементи множин позначатимемо малими літерами латинського алфавіту. Той факт, що об'єкт a є елементом множини M, записують так:

$a \in M$ (читають: "a належить множині M" або "a — елемент множини M"). ***Знак належності*** елемента множині $\in$ — це стилізація першої літери грецького слова εστι (бути). Те, що елемент a не належить множині M, позначають так: $a \notin M$ або $a \bar{\in} M$.

Запис $a, b, c, \ldots \in M$ використовують для скорочення запису $a \in M$, $b \in M$, $c \in M$, … .

Множину називають ***скінченною***, якщо кількість її елементів скінченна, тобто існує натуральне число k, що є кількістю елементів цієї множини. В іншому разі, множина є ***нескінченною***. Кількість елементів скінченної множини A традиційно позначають $|A|$.

Для ***задання множини***, утвореної з будь-яких елементів, будемо використовувати два такі способи. В основі обох із них лежить позначення множини за допомогою фігурних дужок.

1. Якщо $a_1, a_2, \ldots, a_n$ — якісь об'єкти, то їх множину позначають як $\{a_1, a_2, \ldots, a_n\}$, де у фігурних дужках міститься перелік усіх елементів відповідної множини. Так можна задавати тільки скінченні множини. Порядок запису елементів множини в такому позначенні неістотний.

Приклад 1.1. Множину всіх десяткових цифр записують $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, множину основних арифметичних операцій — $\{+, -, \times, /\}$, множину розв'язків нерівності $(x-1)^2 \le 0$ — $\{1\}$.

Одна з основних ідей канторівської теорії множин — розгляд множини як нового самостійного об'єкта математичного дослідження. Тому потрібно розрізняти такі два різні об'єкти, як елемент a та множина $\{a\}$, що складається з єдиного елемента a. Зокрема, множини можуть бути елементами якоїсь іншої множини. Наприклад, множина $D = \{\{a, b\}, \{a, c\}, \{b, c\}\}$ всіх можливих пар з елементів a, b, c складається з трьох елементів, і її задано цілком коректно.

2. Другий спосіб задання множин ґрунтується на зазначенні загальної властивості чи породжувальної процедури для всіх об'єктів, що утворюють описувану множину.

У загальному випадку задання множини M має вигляд $M = \{a \mid P(a)\}$. Цей вираз слід читати так: "M — це множина всіх тих і тільки тих елементів a, для яких виконується умова P". Через $P(a)$ позначено або властивість, яку мають елементи множини M, або якусь породжувальну процедуру, що описує спосіб отримання елементів множини M з уже відомих її елементів чи з інших об'єктів. Замість вертикальної риски іноді пишуть двокрапку.

Приклад 1.2. $S = \{n \mid n$ — непарне число$\}$ або $S = \{n \mid n = 2k + 1, k\in Z\}$, $X = \{x \mid x = \pi k, k\in Z\}$, $F = \{f_i \mid f_{i+2} = f_{i+1} + f_i, i\in N, f_1 = f_2 = 1\}$.

Другий спосіб задання множин більш загальний. Наприклад, уведену вище множину D всіх пар з елементів a, b, c можна задати так:

$$D = \left\{\{x, y\} \mid x \in \{a,b,c\}, y \in \{a,b,c\}, x \neq y\right\}. \qquad (1.1)$$

Для зручності й одностайності виконання математичних викладок уводять поняття множини, яка не містить жодного елемента. Таку множину називають ***порожньою*** та позначають $\varnothing$. Наприклад, якщо досліджують множину об'єктів, які мають задовольняти певні властивості, і з'ясовують, що таких об'єктів не існує, то зручніше сказати, що шукана множина порожня, ніж оголосити її такою, якої немає. Порожню множину можна означати за допомогою будь-якої суперечливої властивості, наприклад, $\varnothing = \{x \mid x \neq x\}$ тощо. Водночас твердженням "множина M непорожня" можна замінювати рівносильне йому твердження "існує елемент, що належать множині M".

1.2. Підмножини

Дві множини A та B називають ***рівними*** (записують $A = B$), якщо вони складаються з тих самих елементів.

Множину A називають ***підмножиною*** множини B (записують $A \subseteq B$ чи $B \supseteq A$) тоді й тільки тоді, коли кожний елемент множини A належить також множині B. Кажуть також, що множина A міститься у множині B. Знаки $\subseteq$ і $\supseteq$ називають ***знаками включення***.

Неважко переконатися, що $A = B$ тоді й тільки тоді, коли одночасно виконуються два включення: $A \subseteq B$ та $B \subseteq A$. Крім того, якщо $A \subseteq B$ та $B \subseteq C$, то $A \subseteq C$. Останні два факти часто використовують у доведеннях тверджень про рівність двох заданих множин.

Якщо $A \subseteq B$, однак $A \neq B$, то пишуть $A \subset B$ і називають множину A ***власною*** (*строгою* або *істинною*) ***підмножиною*** множини B. Знак $\subset$ (або $\supset$), на відміну від знака $\subseteq$ (або $\supseteq$), називається ***знаком строгого включення***.

Очевидно, що для будь-якої множини A виконується включення $A \subseteq A$. Крім того, вважають, що порожня множина є підмножиною будь-якої множини A, тобто $\varnothing \subseteq A$ (зокрема, $\varnothing \subseteq \varnothing$).

Слід чітко розуміти різницю між знаками $\in$ та $\subseteq$ й не плутати ситуації їх використання. Якщо $\{a\} \subseteq M$, то $a\in M$, і навпаки. Однак із включення $\{a\} \subseteq M$, взагалі кажучи, не випливає $\{a\}\in M$. Для будь-якого об'єкта x виконується $x\notin \varnothing$. Наприклад, для множини D (1.1) і її елементів виконуються такі співвідношення:

$\{a, b\}\in D$, $\left\{\{a, b\}, \{b, c\}\right\} \subseteq D$, $a\in \{a, b\}$, $\{c\}\notin \{a, c\}$, $\{a\} \subseteq \{a, b\}$.

Множину всіх підмножин множини A (скінченної чи нескінченної) називають ***булеаном*** множини A і позначають $\beta(A)$. Для булеана множини A використовують також інші позначення: 2^A, $P(A)$, $\mathbf{M}(A)$.

Наприклад, для множини $A = \{a, b\}$ маємо $\beta(A) = \left\{\varnothing, \{a\}, \{b\}, \{a, b\}\right\}$, а для множини $A = \{\varnothing, \{\varnothing\}\}$ $\beta(A) = \left\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\right\}$.

1.1. Задачі і вправи

1. Які з наведених співвідношень правильні?

(а) $\{1, 2, 3\} = \{1, 2, 2, 3\}$; (в) $\{1, 2, 3\} = \{1, 3, 2\}$;

(б) $\{1, 2, 3\} = \left\{1, \{2\}, 3\right\}$; (г) $\{1, 2, 3\} = \left\{\{1, 2\}, \{2, 3\}\right\}$.

2. З яких елементів складається множина B, якщо $A = \{1, 2, 3\}$?

(а) $B = \{y \mid y = x + z, x, z\in A\}$; (в) $B = \{y \mid y = xz, x, z\in A\}$.

(б) $B = \{y \mid x = y + z, x, z\in A\}$;

3. Які з наведених співвідношень правильпі?

(а) $\varnothing = \{0\}$; (в) $\{1, \varnothing\} = \{1\}$; (д) $|\{\varnothing\}| = 0$; (є) $|\{\varnothing, \{\varnothing\}\}| = 2$;

(б) $\varnothing = \{\ \}$; (г) $|\varnothing| = 0$; (е) $|\{\varnothing\}| = 1$; (ж) $|\{\{\varnothing\}\}| = 2$.

4. Які з наведених співвідношень правильні?

(а) $1\in \{1, 2, 3\}$; (е) $\{1, 2\}\in \{1, 2\}$;

(б) $1\in \left\{\{1\}, \{2\}, \{3\}\right\}$; (є) $\{1, 2\}\in \left\{\{1, 2\}\right\}$;

(в) $\{1\}\in \{1, 2, 3\}$; (ж) $\{1, 2\}\in \left\{\{1\}, \{2\}, \{3\}\right\}$;

(г) $\{1\}\in \left\{\{1\}, \{2\}, \{3\}\right\}$; (з) $a\in \{a\}$.

(д) $\{1, 2\}\in \{1, 2, 3\}$;

5. Які з наведених співвідношень правильні?

(а) $0\in \varnothing$; (в) $\varnothing\in \{1\}$; (д) $\varnothing\in \left\{\{\varnothing\}\right\}$;

(б) $\varnothing\in \varnothing$; (г) $\varnothing\in \{\varnothing\}$; (е) $\{\varnothing\}\in \left\{\varnothing, \{\varnothing\}\right\}$.

6. Які з наведених співвідношень правильні?

(а) $1 \subseteq \{1, 2, 3\}$; (г) $\{1\} \subseteq \left\{\{1\}, \{2, 3\}\right\}$;

(б) $\{1\} \subseteq \{1, 2, 3\}$; (д) $\{1, 2\} \subseteq \left\{\{1\}, \{1, 2\}, \{3\}\right\}$;

(в) $\{1, 2, 3\} \subseteq \{1, 2, 3\}$; (е) $\varnothing \subseteq \{1, 2, 3\}$.

7. Нехай $A = \{1, 2, \{1\}\}$. Які з наведених співвідношень правильні?

(а) $1\in A$; (б) $\{1\}\in A$; (в) $\left\{\{1\}\right\}\in A$; (г) $\{1\} \subseteq A$;

(д) $\{\{1\}\} \subseteq A$; (ж) $\{\{2\}\} \subseteq A$; (і) $\varnothing \in A$; (к) $\{\varnothing\} \subseteq A$;
(е) $\{2\} \in A$; (з) $\{1, 2\} \in A$; (ї) $\varnothing \subseteq A$; (л) $\{\varnothing, 1\} \subseteq A$;
(є) $\{2\} \subseteq A$; (и) $\{1, 2\} \subseteq A$; (й) $\{\varnothing\} \in A$; (м) $\{\varnothing, \{1\}\} \subseteq A$.

8. Які з наведених співвідношень правильні?

(а) $\varnothing \subseteq \varnothing$; (г) $\{\varnothing\} \subseteq \{\{\varnothing\}\}$; (є) $\{\{\varnothing\}\} \subseteq \{\varnothing, \{\varnothing\}\}$;
(б) $\varnothing \subseteq \{\varnothing\}$; (д) $\varnothing \subseteq \{1\}$; (ж) $\{\{\varnothing\}\} \subseteq \{\varnothing\}$;
(в) $\{\varnothing\} \subseteq \varnothing$; (е) $\{\varnothing\} \subseteq \{\varnothing\}$; (з) $\{\{\varnothing\}\} \subseteq \{\{\{\varnothing\}\}\}$.

9. Чи існує така одноелементна множина B, що для якоїсь множини A одночасно виконуються співвідношення $A \in B$ та $A \subseteq B$?

10. Для множини A побудувати множину всіх її підмножин, тобто булеан $\beta(A)$:

(а) $A = \{1, 2, 3\}$; (в) $A = \{1, \{2\}, \{1, 2\}\}$;
(б) $A = \{\varnothing\}$; (г) $A = \{\varnothing, \{1, 2\}\}$.

11. Визначити множину:

(а) $\beta(\beta(\{1, 2\}))$; (б) $\beta(\beta(\beta(\varnothing)))$.

1.3. Операції над множинами та їх властивості

Для множин можна ввести теоретико-множинні операції, результатом виконання яких також є множини. За допомогою цих операцій можна конструювати нові множини із заданих.

Нехай A та B — якісь множини.

Об'єднанням множин A та B (позначається $A \cup B$) називається множина тих елементів, які належать хоча б одній із множин A чи B. Символічно операцію об'єднання множин записують так:

$$A \cup B = \{x \mid x \in A \text{ чи } x \in B\}, \text{ або } x \in A \cup B \Leftrightarrow \left[\begin{array}{l} x \in A, \\ x \in B. \end{array}\right.$$

Приклад 1.3. $\{a, b\} \cup \{c, d\} = \{a, b, c, d,\}$, $\{a, c\} \cup \varnothing = \{a, c\}$, $\{a, b, c\} \cup \{a, c, d, e\} = \{a, b, c, d, e\}$.

Перетином множин A та B (позначається $A \cap B$) називається множина, що складається з тих і тільки тих елементів, які належать одночасно множинам A та B, тобто

$$A \cap B = \{x \mid x \in A \text{ і } x \in B\}, \text{ 01> } x \in A \cap B \Leftrightarrow \begin{cases} x \in A, \\ x \in B. \end{cases}$$

Приклад 1.4. $\{a, b, c\} \cap \{a, c, d, e\} = \{a, c\}$, $\{a, b\} \cap \{c, d\} = \varnothing$.

Кажуть, що множини A та B ***не перетинаються***, якщо $A \cap B = \varnothing$.

Операції об'єднання та перетину множин можна поширити на довільну сукупність множин $\{A_i \mid i \in I\}$. Так, об'єднання множин A_i (позначають $\bigcup\limits_{i=I} A_i$) складається з тих елементів, які належать хоча б одній із множин A_i цієї сукупності. А перетин множин A_i (записують $\bigcap\limits_{i=I} A_i$) містить тільки ті елементи, які одночасно належать кожній з множин A_i.

Різницею множин A та B (позначається $A \setminus B$) називається множина тих елементів, які належать множині A й не належать множині B. Отже,

$$A \setminus B = \{x \mid x \in A \text{ і } x \notin B\}, \text{ 01> } x \in A \setminus B \Leftrightarrow \begin{cases} x \in A, \\ x \notin B. \end{cases}$$

Приклад 1.5. $\{b, c\} \setminus \{a, d, c\} = \{b\}$, $\{a, c, d, e\} \setminus \{a, b, c\} = \{d, e\}$, $\{a, b\} \setminus \varnothing = \{a, b\}$, $\{a, b\} \setminus \{a, b, c, d\} = \varnothing$.

Симетричною різницею множин A та B (позначається $A \Delta B$, $A \oplus B$ або $A \div B$) називається множина, що складається з усіх елементів множини A, які не містяться в B, а також усіх елементів множини B, які не містяться в A, тобто

$$A \Delta B = \{x \mid (x \in A \text{ V} x \notin B) \text{ V} (x \in B \text{ V} x \notin A)\}, \text{ 01> } x \in A \Delta B \Leftrightarrow \left[\begin{array}{l} \begin{cases} x \in A, \\ x \notin B, \end{cases} \\ \begin{cases} x \in B, \\ x \notin A. \end{cases} \end{array}\right.$$

Приклад 1.6. $\{a, b, c\} \Delta \{a, c, d, e\} = \{b, d, e\}$, $\{a, b\} \Delta \{a, b\} = \varnothing$, $\{a, b\} \Delta \varnothing = \{a, b\}$.

Уведені теоретико-множинні операції можна проілюструвати так званою ***діаграмою Венна*** (або ***діаграмою Ейлера***) (рис. 1.1).

Тут A та B — це множини точок двох кругів. Тоді множина $A \cup B$ складається з точок областей **I**, **II**, **III**, $A \cap B$ — це область **II**, $A \setminus B$ — область **I**, $B \setminus A$ — область **III**, $A \Delta B$ складається з точок областей **I** та **III**.

У конкретній математичній теорії буває зручно вважати, що всі розглядувані множини є підмножинами певної фіксованої множини, яку називають ***універсальною множиною***, або ***універсумом***, і позначають E (або U). Наприклад, в елементарній алгебрі такою універсальною множиною можна вважати множину дійсних чисел R, у вищій алгебрі — множину комплексних чисел C, в арифметиці — множину цілих чисел Z, у традиційній планіметрії — множину всіх точок площини чи множину всіх геометричних об'єктів, тобто множину множин точок на площині тощо.

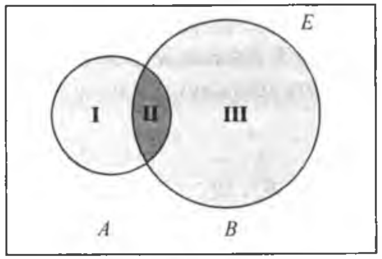


Рис. 1.1

Якщо зафіксовано універсальну множину E, то ***доповненням*** множини A (воно є підмножиною універсальної множини E й позначається $\overline{A}$) називається множина всіх елементів універсальної множини, що не належать множині A, тобто $\overline{A} = \{x \mid x\in E$ та $x\notin A\}$, або $x\in \overline{A} \Leftrightarrow x\notin A$.

Неважко зауважити, що $\overline{A} = E \setminus A$.

Приклад 1.7. Якщо як універсальну множину взяти множину N всіх натуральних чисел, то доповненням P множини P всіх парних натуральних чисел буде множина всіх непарних натуральних чисел.

Зазначимо у вигляді тотожностей *основні властивості* введених вище *теоретико-множинних операцій*:

1) *асоціативність*
$(A \cup B) \cup C = A \cup (B \cup C)$;
$(A \cap B) \cap C = A \cap (B \cap C)$;
2) *комутативність*
$A \cup B = B \cup A$;
$A \cap B = B \cap A$;
3) *дистрибутивність*
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$; (1.2)

4) *ідемпотентність*
$A \cup A = A$;
$A \cap A = A$;
5) *інволютивність*
$\overline{\overline{A}} = A$;
6) *правила (закони) де Моргана*
$\overline{A \cup B} = \overline{A} \cap \overline{B}$;
$\overline{A \cap B} = \overline{A} \cup \overline{B}$.

Правила де Моргана можна узагальнити для сукупності множин:

$$\overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i};\quad \overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i}.$$

Наведемо також інші корисні *теоретико-множинні тотожності*:

$$\begin{array}{llll} A \cup \varnothing = A, & A \cap \varnothing = \varnothing; & A \cup E = E, & A \cap E = A; \\ A \cup \overline{A} = E, & A \cap \overline{A} = \varnothing; & \overline{E} = \varnothing, & \overline{\varnothing} = E. \end{array} \quad (1.3)$$

Окремо запишемо *властивості операції симетричної різниці*:
$A\Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B) = (A \cap \overline{B}) \cup (\overline{A} \cap B)$;
$(A\Delta B)\Delta C = A\Delta(B\Delta C)$ (*асоціативність*);
$A\Delta B = B\Delta A$ (*комутативність*); (1.4)
$A \cap (B\Delta C) = (A \cap B)\Delta(A \cap C)$ (*дистрибутивність перетину*);
$A\Delta A = \varnothing$; $A\Delta E = \overline{A}$; $A\Delta\varnothing = A$.

Приклад 1.8. Покажемо істинність однієї з наведених тотожностей — правила де Моргана

$$\overline{A \cup B} = \overline{A} \cap \overline{B}. \quad (1.5)$$

Доведемо спочатку, що

$$\overline{A \cup B} \subseteq \overline{A} \cap \overline{B}. \quad (1.6)$$

Нехай елемент $x\in \overline{A \cup B}$ Тоді $x\in E \setminus (A \cup B)$, тобто $x\notin A$ та $x\notin B$. Звідси випливає, що $x\in \overline{A}$ і $x\in \overline{B}$; отже, $x\in \overline{A} \cap \overline{B}$. Тому має місце включення (1.6).

Доведемо обернене включення

$$\overline{A} \cap \overline{B} \subseteq \overline{A \cup B}. \quad (1.7)$$

Припустімо, що $x\in \overline{A} \cap \overline{B}$. Це означає, що $x\in \overline{A}$ та $x\in \overline{B}$, тобто $x\notin A$ та $x\notin B$. Тому $x\notin A \cup B$, отже $x\in \overline{A \cup B}$. Зі справедливості обох включень (1.6) і (1.7) випливає істинність рівності (1.5).

Аналогічно можна довести всі інші наведені теоретико-множинні тотожності, які дають змогу спрощувати різні складні вирази з множинами.

Наприклад, послідовно застосовуючи деякі тотожності (1.2) і (1.3), маємо

$$(A \cap B \cap C \cap \overline{D}) \cup (\overline{A} \cap C) \cup (\overline{B} \cap C) \cup (C \cap D) =$$
$$= (A \cap B \cap C \cap \overline{D}) \cup ((\overline{A} \cup \overline{B} \cup D) \cap C) =$$
$$= ((A \cap B \cap D) \cup \overline{A \cap B \cap \overline{D}}) \cap C = E \cap C = C.$$

1.2. Задачі і вправи

1. Нехай $A = \{1, 3, 5, 6\}$, $B = \{1, 2, 3, 5, 7\}$ і $C = \{2, 4, 7\}$. Обчислити

(а) $A \cup B$; (в) $A \cap B \cap C$; (д) $A \Delta B$;
(б) $(A \cup C) \setminus B$; (г) $(A \setminus C) \cup (B \setminus A)$; (е) $(B \setminus C) \cap (A \setminus B)$.

2. За допомогою діаграм Венна перевірити такі теоретико-множинні рівності:

(а) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
(б) $(A \cap B) \cup A = (A \cup B) \cap A = A$;
(в) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$;
(г) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$;
(д) $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$;
(е) $A \Delta B = (A \cup B) \setminus (A \cap B)$.

3. Навести приклад множин A й B, які спростовують рівність

(а) $(A \setminus B) \cup B = A$; (б) $(A \cup B) \setminus B = A$.

Сформулювати та довести необхідні й достатні умови для виконання цієї рівності.

4. Визначити множини

(а) $\varnothing \cap \{\varnothing\}$; (в) $\{\varnothing\} \cup \varnothing$; (д) $\{\varnothing, \{\varnothing\}\} \setminus \{\varnothing\}$;
(б) $\{\varnothing\} \cap \{\varnothing\}$; (г) $\{\varnothing, \{\varnothing\}\} \setminus \varnothing$; (е) $\{\varnothing, \{\varnothing\}\} \setminus \{\{\varnothing\}\}$.

5. Нехай A та B — довільні множини. Довести, що співвідношення, розміщені в одному рядку, рівносильні між собою, тобто зі справедливості одного з них випливає справедливість усіх інших:

(а) $A \subseteq B$, $\overline{B} \subseteq \overline{A}$, $A \cup B = B$, $A \cap B = A$, $A \setminus B = \varnothing$;
(б) $A \subseteq B$, $\overline{A} \cup B = E$, $A \cap \mathrm{B} = \varnothing$;
(в) $A \cap B = \varnothing$, $A \subseteq B$, $B \subseteq A$;
(г) $A \cup B = E$, $A \subseteq B$, $B \subseteq A$.

6. Що можна сказати про множини A та B, якщо

(а) $A \cup B = A \cap B$; (д) $A \setminus B = A$;
(б) $A \setminus B = B \setminus A$; (е) $A \setminus B = \varnothing$;
(в) $A \subseteq \overline{B}$ та $\overline{A} \subseteq B$; (є) $A \setminus B = B$;
(г) $A \cup B = \varnothing$; (ж) $(A \setminus B) \cup (B \setminus A) = A \cup B$?

7. Чи існують множини A, B та C, для яких одночасно виконуються такі співвідношення:

(а) $A \cap B \neq \varnothing$, $A \cap C = \varnothing$, $(A \cap B) \setminus C = \varnothing$;
(б) $A \neq \varnothing$, $A \cap B = \varnothing$, $A \cap C = \varnothing$, $A \setminus (B \cup C) = \varnothing$?

8. Довести тотожності $(A \cap B) \cup A = (A \cup B) \cap A = A$.

9. Сформулювати та довести необхідні й достатні умови для множин A та B, щоб виконувалася рівність $\beta(A) \cup \beta(B) = \beta(A \cup B)$.

10. Що можна сказати про множини A та B, якщо

(а) $A \Delta B = A$; (в) $A \Delta B = \varnothing$; (д) $(A \setminus B) \Delta (B \setminus A) = \varnothing$;
(б) $A \Delta B = \overline{A}$; (г) $A \Delta B = E$; (е) $(A \cup B) \Delta A = B$.

11. Довести, що

(а) $A \cap B \subseteq C \Leftrightarrow A \subseteq \overline{B} \cup C$;
(б) $A \subseteq B \cup C \Leftrightarrow A \cap \overline{B} \subseteq C$.

1.4. Декартів (прямий) добуток множин

Декартовим (прямим) добутком множин A та B (позначається $A \times B$) називається множина всіх пар (a, b), у яких перша компонента належить множині A ($a \in A$), а друга — множині B ($b \in B$), тобто

$$A \times B = \{(a,b) \mid a \in A \text{ і } b \in B\}, \text{ 01> } (a,b) \in A \times B \Leftrightarrow \begin{cases} a \in A, \\ b \in B. \end{cases}$$

Декартів добуток можна природно узагальнити для довільної скінченної сукупності множин. Якщо $A_1, A_2, \ldots, A_n$ — множини, то їх декартовим добутком називається множина

$$D = \{ (a_1, a_2, \ldots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n \},$$

яка складається з усіх наборів $(a_1, a_2, \ldots, a_n)$, у кожному з яких i-й член, що називається ***i-ю координатою***, або ***i-ю компонентою*** набору, належить множині A_i, $i = 1, 2, \ldots, n$. Декартів добуток позначають $A_1 \times A_2 \times \ldots \times A_n$.

Щоб відрізняти набір $(a_1, a_2, \ldots, a_n)$ від множини, що складається з елементів $a_1, a_2, \ldots, a_n$, його записують не у фігурних, а в круглих дужках і називають ***кортежем***, ***вектором*** або ***впорядкованим набором***.

Довжиною кортежу називають кількість його координат.

Два кортежі $(a_1, a_2, \ldots, a_n)$ і $(b_1, b_2, \ldots, b_n)$ однакової довжини вважають ***рівними*** тоді й тільки тоді, коли рівні відповідні їх координати, тобто $a_i = b_i$, $i = 1, 2, \ldots, n$. Отже, кортежі (a, b, c) й (a, c, b) різні, а множини $\{a, b, c\}$ й $\{a, c, b\}$ — рівні між собою.

Декартів добуток множини A на себе n разів, тобто множину $A \times A \times \ldots \times A$ називають ***n*-м декартовим** (**прямим**) **степенем** множини A та позначають A^n. Вважають, що $A^0 = \varnothing$ $(n = 0)$ й $A^1 = A$ $(n = 1)$.

Приклад 1.9. 1. Якщо $A = \{a, b\}$ та $B = \{b, c, d\}$, то
$A \times B = \{(a, b), (a, c), (a, d), (b, b), (b, c), (b, d)\}$,
$A^2 = \{(a, a), (a, b), (b, a), (b, b)\}$.

2. Якщо R — множина дійсних чисел, або множина точок координатної прямої, то R^2 — це множина пар (a, b), де $a, b \in R$, або множина точок координатної площини.

Координатне зображення точок площини вперше запропонував французький математик і філософ Рене Декарт, тому введена теоретико-множинна операція й називається декартовим добутком.

3. Скінченна множина A, елементами якої є символи (літери, цифри, спеціальні знаки тощо), називається ***алфавітом***. Елементи декартового степеня A^n називають ***словами*** довжини n в алфавіті A.

$$A^* = \{e\} \cup A \cup A^2 \cup A^3 \cup \ldots = \{e\} \cup \bigcup_{i \in N} A^i \quad —$$

це *множина всіх слів у алфавіті* A; тут e — *порожнє слово* (слово довжини 0), яке не містить жодного символу алфавіту A. Порожнє слово позначають також через ε.

Замість запису слів із прямого степеня A^n у вигляді кортежів $(a_1, a_2, \ldots, a_n)$ частіше використовують традиційну форму їх запису у вигляді послідовності символів $a_1a_2\ldots a_n$, $a_j \in A$, $j = 1, 2, \ldots, n$. Наприклад, 010111, 011, 0010, 100 — слова в алфавіті $B = \{0, 1\}$, а 67–35, 98)) * (+ 1, (450 + 12)/27, 349 * 2 + 17 — це слова в алфавіті $C = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, –, *, /, (,)\}$.

Операція декартового добутку неасоціативна й некомутативна, тобто множини $(A \times B) \times C$, $A \times (B \times C)$ й $A \times B \times C$, а також множини $A \times B$ та $B \times A$, узагалі кажучи, різні. (Зауважимо, що в математиці множини $(A \times B) \times C$ й $A \times (B \times C)$ іноді ототожнюють з $A \times B \times C$, вважаючи, що кортежі $((a, b), c)$, $(a, (b, c))$ й (a, b, c) збігаються [19, 26].)

Зв'язок декартового добутку з іншими теоретико-множинними операціями виражають такі тотожності:

$$\begin{aligned} (A \cup B) \times C &= (A \times C) \cup (B \times C), \\ (A \cap B) \times C &= (A \times C) \cap (B \times C), \\ A \times (B \cup C) &= (A \times B) \cup (A \times C), \\ A \times (B \cap C) &= (A \times B) \cap (A \times C). \end{aligned} \tag{1.8}$$

Проекцією на і-ту вісь (***і-ю проекцією***) кортежу $w = (a_1, a_2, \ldots, a_n)$ називається i-та координата a_i кортежу w; позначається $\mathrm{Pr}_i w$.

Проекцією кортежу $w = (a_1, a_2, \ldots, a_n)$ на осі з номерами $i_1, i_2, \ldots, i_k$ називається кортеж $(a_{i_1}, a_{i_2}, \ldots, a_{i_k})$; позначається $\mathrm{Pr}_{i_1, i_2, \ldots, i_k} w$.

Нехай V — множина кортежів однакової довжини. Проекцією множини V на i-ту вісь (позначається $\mathrm{Pr}_i V$) називається множина проекцій на i-ту вісь усіх кортежів множини V: $\mathrm{Pr}_i V = \{\mathrm{Pr}_i v \mid v \in V\}$.

Аналогічно означають проекцію множини V на кілька осей:

$$\mathrm{Pr}_{i_1, i_2, \ldots, i_k} V = \{\mathrm{Pr}_{i_1, i_2, \ldots, i_k} v \mid v \in V\}.$$

Приклад 1.10. $\mathrm{Pr}_{i_1, i_2, \ldots, i_k}(A_1 \times A_2 \times \ldots \times A_n) = A_{i_1} \times A_{i_2} \times \ldots \times A_{i_k}$. Якщо $V = \{(a, b, c), (a, c, d), (a, b, d)\}$, то $\mathrm{Pr}_1 V = \{a\}$, $\mathrm{Pr}_2 V = \{b, c\}$, $\mathrm{Pr}_{1,3} V = \{(a, c), (a, d)\}$, $\mathrm{Pr}_{2,3} V = \{(b, c), (c, d), (b, d)\}$.

1.3. Задачі і вправи

1. Для заданих множин $A = \{1, 2\}$ та $B = \{2, 3, 4\}$ визначити
(а) $A \times B$; (в) B^2; (д) $A \times B \times A$;
(б) $B \times A$; (г) $(B \setminus A) \times A$; (е) $A \times (A \cup B)$.

2. Довести, що існують множини A, B та C такі, що
(а) $A \times B \neq B \times A$; (б) $(A \times B) \times C \neq A \times (B \times C)$.
Для яких множин виконуються рівності?

3. Довести, що
(а) $(A \times B) \cap (B \times A) = (A \cap B) \times (A \cap B)$;
(б) $(A \times B) \cup (B \times A) \subseteq (A \cup B) \times (A \cup B)$.

4. Сформулювати та довести необхідні й достатні умови виконання рівності $(A \cup B) \times (A \cup B) = (A \times B) \cup (B \times A)$.

5. Позначимо $D = \beta(M) \times \beta(M) \times \beta(M)$, де $M = \{1, 2\}$. Виписати всі такі кортежі $(A, B, C) \in D$, що
(а) $A \cap B = C$; (б) $A \cup B \cup C = M$; (в) $A \cap B \neq B \cap C$.

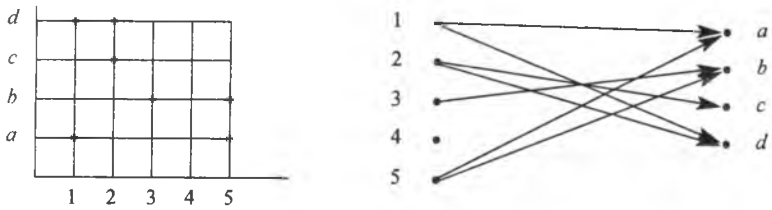
6. Довести, що коли $B \neq \varnothing$, то
(а) $\mathrm{Pr}_1(A \times B) = A$;
(б) якщо $C \subseteq A \times B$, то $\mathrm{Pr}_1 C \subseteq A$.

1.5. Відповідності

Відповідністю між множинами A та B називається будь-яка підмножина $C \subseteq A \times B$. Якщо $(a, b) \in C$, то кажуть, що ***елемент*** b ***відповідає елементу*** a у відповідності C.

Оскільки відповідності — це множини, то для їх задання використовують ті самі методи, що й для довільних множин.

Крім того, відповідність можна задавати (чи ілюструвати) за допомогою так званого ***графіка відповідності***. Нехай $A = \{1, 2, 3, 4, 5\}$ і $B = \{a, b, c, d\}$, а $C = \{(1, a), (1, d), (2, c), (2, d), (3, b), (5, a), (5, b)\}$ — відповідність між A та B. Позначимо числами 1, 2, 3, 4, 5 вертикальні прямі, а літерами a, b, c, d — горизонтальні прямі на координатній площині (рис. 1.2, *а*). Тоді виділені вузли на перетині цих прямих позначають елементи відповідності C та утворюють графік відповідності C.



а *б*

Рис. 1.2

Зручним методом задання невеликих скінченних відповідностей є ***діаграма***, або ***граф відповідності***. В одній колонці розміщують точки, позначені елементами множини A, у другій (праворуч) — точки, позначені елементами множини B. Із точки a першої колонки проводять стрілку в точку b другої колонки тоді й тільки тоді, коли пара (a, b) належить заданій відповідності. На рис. 1.2, *б* зображено діа-граму відповідності C з попереднього абзацу.

Відповідність можна задавати, означаючи співвідношення, які мають задовольняти обидві її координати. Наприклад, розглянемо на класичній координатній площині $R^2 = R \times R$ такі відповідності: $C_1 = \{(x, y) \mid x^2 + y^2 = 1\}$, $C_2 = \{(x, y) \mid y = x^2\}$, $C_3 = \{(x, y) \mid |x| \le 1, |y| \le 1\}$. Графік відповідності C_1 — це коло радіуса 1 із центром у початку координат, графік C_2 — квадратична парабола, а графік C_3 — усі точки квадрата з вершинами $(-1, -1)$, $(-1, 1)$, $(1, 1)$ і $(1, -1)$.

Нехай $C \subseteq A \times B$ — відповідність між множинами A і B.

Множина $\mathrm{Pr}_1 C$ називається ***областю визначення***, а множина $\mathrm{Pr}_2 C$ — ***областю значень*** відповідності C (інші позначення — відповідно δ_C та ρ_C).

Образом елемента $a \in \mathrm{Pr}_1 C$ при відповідності C називається множина всіх елементів $b \in \mathrm{Pr}_2 C$, що відповідають елементу a; позначається $C(a)$. ***Прообразом елемента*** $b \in \mathrm{Pr}_2 C$ при відповідності C називається множина всіх тих елементів $a \in \mathrm{Pr}_1 C$, яким відповідає елемент b; позначається $C^{-1}(b)$. Якщо $D \subseteq \mathrm{Pr}_1 C$, то ***образом множини*** D при відповідності C називається об'єднання образів усіх елементів з D; позначається $C(D)$. Аналогічно означають ***прообраз множини*** $G \subseteq \mathrm{Pr}_2 C$; позначають $C^{-1}(G)$.

Оскільки відповідності — це множини, то до них можна застосовувати всі відомі теоретико-множинні операції: об'єднання, перетин, різницю тощо. Зокрема, для операції доповнення відповідності C між множинами A і B універсальною множиною є $A \times B$.

Додатково введемо для відповідностей дві специфічні операції.

Відповідністю, ***оберненою*** до заданої відповідності C між множинами A та B, називається така відповідність D між множинами B і A, що $D = \{(b, a) \mid (a, b) \in C\}$. Відповідність, обернену до відповідності C, позначають C^{-1}.

Якщо задано відповідності $C \subseteq A \times B$ та $D \subseteq B \times F$, то ***композицією*** (***суперпозицією***, ***добутком***) ***відповідностей*** C і D (позначається $C \circ D$) називається така відповідність H між множинами A і F:

$$H = \{(a, b) \mid \text{існує елемент } c \in B, \text{ для якого } (a, c) \in C \text{ та } (c, b) \in D\}.$$

Розглянемо окремі важливі випадки відповідностей C між множинами A і B.

Якщо $\mathrm{Pr}_1 C = A$, то відповідність C називається ***всюди*** (***скрізь***) ***визначеною***, а в іншому разі — ***частковою***.

Відповідність $f \subseteq A \times B$ називається ***функціональною***, або ***функцією*** з A в B, якщо кожному елементу $a \in \mathrm{Pr}_1 f$ відповідає тільки один елемент із $\mathrm{Pr}_2 f$, тобто образом кожного елемента $a \in \mathrm{Pr}_1 f$ є єдиний елемент b з $\mathrm{Pr}_2 f$. Якщо f — функція з A в B, то кажуть, що функція має ***тип*** $A \to B$ і позначають $f: A \to B$ чи $A \xrightarrow{f} B$.

Усюди визначена функціональна відповідність $f \subseteq A \times B$ називається ***відображенням*** з A в B; позначається як і функція $f: A \to B$ або $A \xrightarrow{f} B$. Відображення називають також *всюди (скрізь) визначеними функціями*.

Відображення типу $A \to A$ називають ***перетворенням*** множини A.

Через B^A позначають множину всіх відображень з A в B.

Оскільки функція і відображення є окремими випадками відповідності, то для них мають місце всі наведені вище означення: поняття областей визначення та значень, образу та прообразу елементів і множин тощо. Зокрема, для функції f елементи множини $\mathrm{Pr}_1 f$ називають ***аргументами*** функції, образ $f(a)$ елемента $a \in \mathrm{Pr}_1 f$ — ***значенням*** функції f на a.

Відповідність C називається ***сюр'єктивною*** (***сюр'єкцією***), або ***відповідністю на*** множину B, якщо $\mathrm{Pr}_2 C = B$.

Відповідність C називається ***ін'єктивною*** (***ін'єкцією***), або ***різнозначною відповідністю***, якщо для кожного елемента $b \in \mathrm{Pr}_2 C$ його прообраз $C^{-1}(b)$ складається тільки з одного елемента. Інакше кажучи, різним елементам множини A відповідають різні елементи множини B. Іноді ін'єкцію називають 1–1 відповідністю.

Відображення, яке є водночас сюр'єктивним й ін'єктивним, називається ***бієктивним***, або ***бієкцією***. Бієктивні відображення називають часто також ***взаємно однозначними відображеннями*** чи ***взаємно однозначними відповідностями*** між множинами A і B.

Отже, відповідність взаємно однозначна тоді й лише тоді, коли вона функціональна, всюди визначена, сюр'єктивна та ін'єктивна.

Відповідність $i_A = \{(a, a) \mid a \in A\}$ називають ***тотожним перетворенням, діагональною відповідністю***, або ***діагоналлю*** в A.

Бієктивне відображення з A в A називають ***підстановкою*** множини A.

Для довільної відповідності C між A і A позначимо через $C^{(n)}$ відповідність $C \circ C \circ \ldots \circ C$ (n входжень літери C). Вважають, що $C^{(0)} = i_A$ та $C^{(1)} = C$.

Наведемо приклади відповідностей, відображень і функцій.

Приклад 1.11. 1. Відповідність між клітинками та фігурами на шахівниці в будь-який момент гри функціональна, але це не відображення, бо не всі поля шахівниці зайнято фігурами.

2. Відповідність між натуральними числами і сумами цифр їх десяткового запису є відображенням. Воно не є ін'єктивним, бо йому належать, наприклад, такі пари (17, 8) і (26, 8).

3. Відповідність, за якою кожному натуральному числу n відповідає число $3n$, очевидно, є взаємно однозначною відповідністю між множиною всіх натуральних чисел і множиною натуральних чисел, кратних 3.

4. Відповідність між множиною точок координатної площини R^2 і множиною всіх векторів із початком у точці (0, 0) є взаємно однозначною.

1.4. Задачі і вправи

1. Нехай задано множини $A = \{a, b, c, d\}$ та $B = \{1, 2, 3, 4, 5\}$ і відповідності між A і B:

$C_1 = \{(a, 1), (a, 3), (a, 5), (b, 1), (b, 3), (d, 3), (d, 4), (d, 5)\}$,
$C_2 = \{(a, 4), (a, 5), (b, 2), (b, 3), (c, 1), (d, 2), (d, 3)\}$,
$C_3 = \{(b, 1), (b, 2), (b, 3), (c, 1), (c, 2), (d, 1), (d, 5)\}$.

Визначити

(а) $\mathrm{Pr}_j C_i$, $j = 1, 2$, $i = 1, 2, 3$;
(б) $C_1 \cup C_2$, $C_2 \cup C_3$;
(в) $C_1 \cap C_2$, $C_2 \cap (C_1 \cup C_3)$;
(г) $C_1 \setminus C_2$, $C_3 \setminus (C_1 \cap C_3)$;
(д) $\complement_1$, $C_2 \Delta C_3$;
(е) C_i^{-1}, $i = 1, 2, 3$.

Побудувати графіки й діаграми відповідностей C_1, C_2, C_3 та відповідностей з пунктів (б)–(е).

2. Нехай задано множини $A = \{a, b, c, d\}$, $B = \{1, 2, 3, 4, 5\}$ і $G = \{\alpha, \beta, \gamma\}$, відповідності між A та B

$C_1 = \{(a, 2), (a, 3), (b, 1), (c, 4), (c, 5), (d, 2), (d, 3)\}$,
$C_2 = \{(b, 3), (b, 5), (c, 1), (c, 4), (d, 1), (d, 4), (d, 5)\}$

і відповідності між B і G

$D_1 = \{(1, \beta), (1, \gamma), (2, \alpha), (2, \gamma), (3, \gamma), (5, \alpha), (5, \beta)\}$,
$D_2 = \{(1, \alpha), (2, \alpha), (2, \beta), (3, \alpha), (4, \alpha), (4, \gamma)\}$.

Визначити: (а) $C_i \circ D_j$, $i, j = 1, 2$; (б) $C_i \circ C_j^{-1}$, $i, j = 1, 2$; (в) $C_2 \circ (D_1 \circ D_2^{-1})$.

3. Нехай задано такі відповідності між R і R:

$C_1 = \{(x, y) \mid x^2 + y^2 \geq 1\}$;
$C_2 = \{(x, y) \mid |x| + |y| \geq 3\}$;
$C_3 = \{(x, y) \mid y^2 - |x| \geq 0\}$;
$C_4 = \{(x, y) \mid (|x| + 0{,}5)^{|x| + |y|} \leq x^2 + |x| + 0{,}25\}$.

Побудувати графік відповідності:

(а) $\overline{C_1 \cap C_2} \cap C_3$;
(б) $C_3 \cup C_4$;
(в) $C_1 \cap C_2 \cap C_4$.

4. Нехай $C = \{(x, y) \mid x^2 + y^2 = 1\}$ — відповідність між R і R. Для $n \in N$ визначити відповідність $C^{(n)}$.

5. Нехай задано відповідності між R і R: $L_1 = \{(x, y) \mid y = ax + b\}$ і $L_2 = \{(x, y) \mid y = cx + d\}$. Для яких значень параметрів $a, b, c, d \in R$ виконується рівність $L_1 \circ L_2 = L_2 \circ L_1$?

6. Що можна сказати про множини A і B, якщо

(а) $i_A \subseteq A \times B$;
(б) $i_A \cap (A \times B) = \varnothing$;
(в) з $(a, b) \in A \times B$ випливає $(b, a) \in A \times B$;
(г) з $(a, b) \in A \times B$ випливає $a \neq b$;
(д) $A \times B \subseteq B \times A$;
(е) $|A \times B| = 1$?

7. Нехай C_1 — відповідність між A і B, C_2 — відповідність між B і G. Довести такі твердження:

(а) якщо $C_1 \circ C_2 = D$, то $\Pr_1 D \subseteq \Pr_1 C_1$ і $\Pr_2 D \subseteq \Pr_2 C_2$;

(б) $C_1 \circ C_2 = \varnothing \Leftrightarrow \Pr_2 C_1 \cap \Pr_1 C_2 = \varnothing$.

8. Що можна сказати про відповідність C між множинами A і B, якщо

(а) для будь-якого $x \in A$ існує $y \in B$ такий, що $(x, y) \in C$;

(б) з $(x, y), (x, z) \in C$ випливає $y = z$;

(в) з $(x, y), (z, y) \in C$ випливає $x = z$;

(г) для будь-якого $y \in B$ існує $x \in A$ такий, що $(x, y) \in C$;

(д) для будь-якого $x \in A$ існує єдиний елемент $y \in B$ такий, що $(x, y) \in C$, і навпаки, для будь-якого $t \in B$ існує єдиний елемент $z \in A$ такий, що $(z, t) \in C$?

9. Нехай задано такі відповідності між множинами $A = \{a, b, c, d, e\}$ та $B = \{1, 2, 3, 4, 5\}$:

$C_1 = \{(a, 2), (a, 5), (b, 1), (b, 5), (c, 2), (d, 3), (d, 5)\}$,

$C_2 = \{(a, 3), (b, 2), (c, 3), (e, 3)\}$,

$C_3 = \{(a, 2), (b, 3), (c, 4), (d, 1), (e, 5)\}$,

$C_4 = \{(a, 2), (a, 3), (b, 1), (c, 4), (c, 5), (d, 1), (e, 2), (e, 4)\}$,

$C_5 = \{(a, 4), (b, 3), (c, 5), (d, 2), (e, 1)\}$.

Визначити, які з цих відповідностей

(а) всюди визначені; (г) сюр'єктивні;

(б) функціональні; (д) бієктивні (взаємно однозначні).

(в) ін'єктивні;

Побудувати графіки та діаграми цих відповідностей.

10. Довести, що відповідність f між A та B є відображенням тоді й тільки тоді, коли $i_A \subseteq f \circ f^{-1}$ і $f^{-1} \circ f \subseteq i_B$.

11. Довести, що об'єднання $f_1 \cup f_2$ двох відображень f_1 і f_2 з A в B буде відображенням з A в B тоді й тільки тоді, коли $f_1 = f_2$.

12. Встановити взаємно однозначну відповідність між множинами

(а) A^k і A^{N_k}; (б) $\beta(A)$ і $\{0, 1\}^A$.

13. Нехай f — відповідність між A і B. Довести, що:

(а) f всюди визначена тоді й тільки тоді, коли $i_A \subseteq f \circ f^{-1}$;

(б) f функціональна тоді й тільки тоді, коли $f^{-1} \circ f \subseteq i_B$;

(в) f сюр'єктивна тоді й тільки тоді, коли $i_B \subseteq f^{-1} \circ f$;

(г) f ін'єктивна тоді й тільки тоді, коли $f \circ f^{-1} \subseteq i_A$;

(д) f взаємно однозначна тоді й тільки тоді, коли $f \circ f^{-1} = i_A$ і $f^{-1} \circ f = i_B$.

1.6. Рівнопотужність множин

Усі введені вище теоретико-множинні операції та їх властивості мають місце як для скінченних, так і для нескінченних множин. Істотна різниця між цими множинами виявляється тоді, коли мова за-

ходить про "кількість елементів" і в разі спроби порівняти множини за цією ознакою. Тут слова "кількість елементів" узято в лапки тому, що зрозуміла умовність і невизначеність цього поняття для нескінченних множин.

Одні з основних досягнень канторівської теорії множин — це поширення поняття "кількість елементів" зі скінченних множин на нескінченні та формулювання принципу, згідно з яким можна порівнювати нескінченні множини за "кількістю елементів". Зокрема, не-сподіваним і незвичайним виявився той факт, що різні нескінченні множини можуть мати різну "кількість елементів", тобто для нескінченностей також існує своя ієрархія.

Канторівська ідея ґрунтується на такому простому спостереженні: щоб порівняти за кількістю елементів дві скінченні множини, зовсім не обов'язково перелічувати елементи кожної з них. Можна діяти так. Наприклад, нехай потрібно порівняти за кількістю елементів дві множини — множину S студентів і множину M усіх місць в аудиторіях факультету. Запропонуємо кожному студентові зайняти одне місце. Якщо кожен студент отримає місце і в аудиторіях не залишиться жодного вільного місця, то кількість елементів у множинах S і M однакова. В іншому разі множина S містить більше елементів, ніж множина M, або навпаки. Запропонована процедура встановлює певну функціональну відповідність між множинами S і M. У першому випадку ця відповідність виявляється взаємно однозначною, а у другому та третьому умови взаємної однозначності не виконуються: або порушується умова повної визначеності (принаймні один студент не дістав місця), або порушується умова сюр'єктивності (хоча б одне місце залишилося вільним).

Принагідно зауважимо, що багато хто з математиків вважає, що описаний простий спосіб порівняння кількостей елементів у двох скінченних множинах логічно передує виникненню поняття числа.

Нагадаємо, що кількість елементів скінченної множини A позначають через $|A|$.

Отже, неважко переконатися, що між двома скінченними множинами A та B існує взаємно однозначна відповідність тоді й тільки тоді, коли $|A| = |B|$.

Сформульоване твердження дає змогу розв'язувати задачу обчислення кількості елементів множини A шляхом встановлення взаємно однозначної відповідності між множиною A та якоюсь множиною B, кількість елементів якої відома або легко може бути визначена. Для ілюстрації цього методу доведемо важливу теорему про кількість підмножин заданої скінченної множини.

Теорема 1.1. Нехай $A = \{a_1, a_2, \ldots, a_n\}$ — скінченна множина з n елементів ($|A| = n$). Тоді кількість усіх її підмножин дорівнює 2^n, тобто $2^{|A|}$.

Доведення. Розглянемо множину всіх кортежів $(b_1, b_2, \ldots, b_n)$ довжини n, що складаються з двійкових цифр 0 або 1 (тобто $b_i \in B = \{0, 1\}$, $i = 1, 2, \ldots, n$). Множина таких кортежів — це B^n.

Установимо відповідність між підмножинами множини A та кортежами з B^n. Кожній підмножині $A' \subseteq A$ поставимо у відповідність такий двійковий кортеж $(b_1, b_2, \ldots, b_n)$, у якому $b_i = 1$, якщо $a_i \in A'$, і $b_i = 0$, якщо $a_i \notin A'$, $i = 1, 2, \ldots, n$.

За цим правилом порожній множині $\varnothing \subseteq A$ відповідатиме кортеж $(0, 0, \ldots, 0)$, самій множині A — кортеж $(1, 1, \ldots, 1)$, а підмножині $A' = \{a_i, a_j\}$, $i, j = 1, 2, ..., n$, $i < j$, — кортеж $(0,..., 0, 1, 0, \ldots, 0, 1, 0, \ldots,0)$, у якому двійкові цифри 1 стоять на i-му та j-му місцях. Установлена відповідність взаємно однозначна. Отже, кількість усіх підмножин множини A дорівнює $|B^n|$.

Методом математичної індукції доведемо, що $|B^n| = 2^n$.

Для $n = 1$ маємо $B^1 = B$ і $|B| = 2 = 2^1$.

Припустімо, що $|B^{k-1}| = 2^{k-1}$. Із того, що кожному елементу $(b_1, b_2, \ldots, b_{k-1})$ множини B^{k-1} відповідають два елементи $(b_1, b_2, \ldots, b_{k-1}, 0)$ і $(b_1, b_2, \ldots, b_{k-1}, 1)$ множини B^k, випливає, що кількість елементів множини B^k вдвічі більша, ніж кількість елементів множини B^{k-1}, тобто $|B^k| = |B^{k-1}| \cdot 2 = 2^{k-1} \cdot 2 = 2^k$. Теорему 1.1 доведено.

Отже, з доведеної теореми випливає, що для скінченної множини A виконується рівність $|\beta(A)| = 2^{|A|}$.

Назвемо множини A і B ***рівнопотужними***, або множинами, які мають рівні (однакові) потужності, якщо існує взаємно однозначна відповідність між A і B; рівнопотужність множин A і B позначають $A \sim B$.

Отже, дві скінченні множини A та B мають однакову потужність тоді й лише тоді, коли вони складаються з однакової кількості елементів. Тому поняття потужності є узагальненням поняття кількості елементів множини.

Звернімо увагу на те, що ми не означили безпосередньо поняття “потужність множини”, а лише дали означення рівнопотужності множин. Кантор пропонував розуміти під потужністю ту спільну властивість, яку мають усі рівнопотужні множини. Виходячи з того, що для рівнопотужних скінченних множин такою спільною властивістю є кількість їх елементів, за аналогією переносять цю властивість на нескінченні множини, хоча це, взагалі кажучи, не зовсім коректно, у чому ми переконаємося нижче.

Неважко довести такі *властивості рівнопотужності*:

1) $A \sim A$ (*рефлексивність*);
2) якщо $A \sim B$, то $B \sim A$ (*симетричність*); (1.9)
3) якщо $A \sim B$ і $B \sim C$, то $A \sim C$ (*транзитивність*).

Наведемо кілька прикладів рівнопотужних нескінченних множин.

Приклад 1.12. 1. Множина натуральних чисел N рівнопотужна множині $S = \{1, 4, 9, 16, \ldots\}$, що складається з квадратів натуральних чисел. Потрібну взаємну однозначну відповідність можна задати законом (n, n^2), $n \in N$, $n^2 \in S$.

2. Множина Z усіх цілих чисел рівнопотужна множині P всіх парних чисел. У цьому разі взаємно однозначну відповідність можна задати так: $(n, 2n)$, $n \in Z$, $2n \in P$.

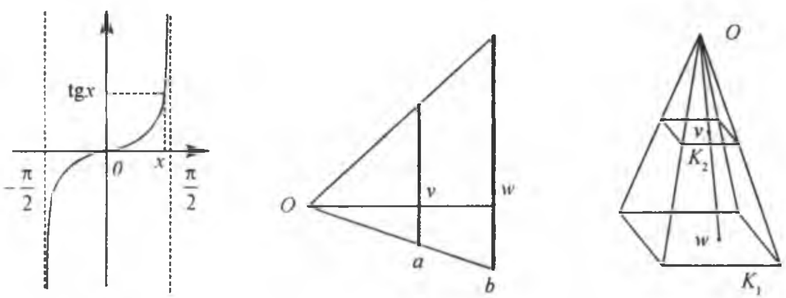
3. Множина точок інтервалу $(-\pi/2, \pi/2)$ рівнопотужна множині точок дійсної прямої. Шукану взаємно однозначну відповідність можна задати, наприклад, за допомогою тригонометричної функції tg: $(x, \operatorname{tg} x)$, $x \in (-\pi/2, \pi/2)$, $\operatorname{tg} x \in (-\infty, \infty)$ (рис. 1.3, *а*).

4. Множини точок двох довільних відрізків a і b рівнопотужні. Правило, за яким можна задати взаємно однозначну відповідність між точками відрізків a і b різної довжини, проілюстровано на рис. 1.3, *б*. Кожний промінь із початком у точці O, що перетинає відрізки a і b в точках v та w, утворює одну пару (v, w) потрібної взаємно однозначної відповідності.

5. Аналогічно можна задати взаємно однозначну відповідність між множинами точок двох довільних квадратів K_1 і K_2 різних розмірів (рис. 1.3, *в*).

Зауваження. Із рівнопотужності довільних відрізків і транзитивності рівнопотужності випливає, що будь-який відрізок рівнопотужний інтервалу $(-\pi/2, \pi/2)$, а значить, і всій прямій.

З усіх наведених прикладів випливає, що нескінченна множина може бути рівнопотужна своїй власній підмножині, що неможливо для скінченних множин. Саме незвичність і екзотичність висновків типу “множина парних чисел містить стільки ж елементів, як і множина всіх цілих чисел", "будь-який інтервал містить стільки ж точок, як і вся пряма" тощо призвели до того, що в канторівської теорії множин поряд із палкими прихильниками було чимало рішучих противників. Вони категорично відкидали всі спроби досліджувати та порівнювати нескінченні множини, зокрема, на тій підставі, що "частина завжди "менша" від цілого й не може бути "рівна" йому". Але це не злякало Кантора. Він зрозумів і своїми результатами переконував інших, що для нескінченних множин виконуються інші закони, непридатні для скінченних множин. Розвиваючи цю тезу, Р. Дедекінд узагалі запропонував вважати нескінченною таку множину, яка рівнопотужна своїй власній підмножині, тобто покласти цю "дивну" властивість в основу означення нескінченної множини.



а б в

Рис. 1.3

Наступне питання, що постало перед Кантором: чи всі нескінченні множини рівнопотужні?

1.5. Задачі і вправи

1. Довести такі властивості рівнопотужності множин:

(а) $A \sim A$ (рефлексивність);

(б) якщо $A \sim B$, то $B \sim A$ (симетричність);

(в) якщо $A \sim B$ і $B \sim C$, то $A \sim C$ (транзитивність);

2. Встановити взаємно однозначну відповідність:

(а) між множиною непарних натуральних чисел і множиною всіх натуральних чисел N;

(б) між множиною натуральних чисел N і множиною цілих чисел Z;

(в) між множиною натуральних чисел N і множиною цілих чисел, кратних числу k, $k \in N$.

3. Чи правильні такі твердження:

(а) якщо $A = B$, то $A \sim B$;

(б) якщо $A \sim B$, то $A = B$?

4. Довести, що:

(а) скінченна множина не рівнопотужна жодній своїй власній підмножині;

(б) скінченна множина не рівнопотужна жодній множині, яка містить її як власну підмножину.

5. Довести, що для довільних множин A_1 і A_2 існують такі множини B_1 і B_2, що $A_1 \sim B_1$, $A_2 \sim B_2$ та $B_1 \cap B_2 = \varnothing$.

6. Довести, що коли $A \sim B$ та $C \sim D$, то $A \times C \sim B \times D$.

1.7. Зліченні множини

Множина A, рівнопотужна множині N натуральних чисел, називається ***зліченною***.

Інакше кажучи, зліченна множина A — це така множина, всі елементи якої можна занумерувати числами 1, 2, 3, ..., тобто можна зазначити спосіб, за яким першому елементу множини A поставлено у відповідність число 1, другому — число 2, третьому — число 3 і т. д. Отже, будь-яку зліченну множину A можна подати у вигляді $A = \{a_1, a_2, a_3, \ldots, a_n, \ldots\}$.

Неважко переконатися, що множини квадратів натуральних чисел, усіх парних чисел, усіх непарних чисел, чисел, кратних якомусь числу k, чисел, які закінчуються парою цифр 00, тощо зліченні.

Перейдемо до вивчення властивостей зліченних множин.

Теорема 1.2. Будь-яка нескінченна множина M містить зліченну підмножину.

Доведення. Оскільки M — нескінченна множина, візьмемо два елементи $a_1, b_1 \in M$ $(a_1 \neq b_1)$. Очевидно, що множина $M \setminus \{a_1, b_1\}$ нескінченна. Тоді візьмемо наступні два нові елементи $a_2, b_2 \in M \setminus \{a_1, b_1\}$ $(a_2 \neq b_2)$ і т. д. Отже, можна виділити з множини M дві зліченні підмножини $A = \{a_1, a_2, \ldots, a_n, \ldots\} \subseteq M$ і $B = \{b_1, b_2, \ldots, b_n, \ldots\} \subseteq M$. Це дає змогу підсилити формулювання теореми, а саме: будь-яка нескінченна множина M містить таку зліченну підмножину A, що множина $M \setminus A$ є нескінченною (бо $B \subseteq M \setminus A$).

Теорема 1.3. Будь-яка підмножина зліченної множини або скінченна, або зліченна.

Доведення. Нехай $A = \{a_1, a_2, \dots, a_n, \dots\}$ — зліченна множина, і $B \subseteq A$. Отже, $B = \{a_{i_1}, a_{i_2}, \dots, a_{i_k}, \dots\}$, і можливі дві ситуації: або остання послідовність уривається на якомусь елементі (тоді B — скінченна множина), або вона нескінченна, і для неї, задавши відповідність (l, a_{i_l}), $l \in N$, одержимо, що B — зліченна множина.

Із теорем 1.2 та 1.3, зокрема, випливає, що зліченні множини — це певною мірою найпростіші нескінченні множини, бо, з одного боку, вони містяться в будь-якій нескінченній множині, а з іншого — містять у собі тільки скінченні множини або нескінченні множини, які є зліченними.

Теорема 1.4. Об'єднання скінченної чи зліченної сукупності скінченних або зліченних множин є скінченною чи зліченною множиною.

Доведення. Кожну з множин A_i даної сукупності можна подати у вигляді $A_i = \{a_1^i, a_2^i, \dots, a_n^i, \dots\}$. Остання послідовність нескінченна, коли A_i — зліченна множина, або ж ця послідовність закінчується якимось елементом a_n^i, якщо A_i — скінченна множина.

Усі елементи множин A_i запишемо у вигляді такої таблиці (зі скінченною чи зліченною множиною рядків):

$$\begin{array}{lcl}
a_1^1,\ a_2^1,\ a_3^1,\ \dots,\ a_n^1,\ \dots & & A_1 \\
\ \nearrow\ \ \nearrow\ \ \nearrow & & \\
a_1^2,\ a_2^2,\ a_3^2,\ \dots,\ a_n^2,\ \dots & & A_2 \\
\ \nearrow\ \ \nearrow & & \\
a_1^3,\ a_2^3,\ a_3^3,\ \dots,\ a_n^3,\ \dots & & A_3 \\
\ \nearrow & & \\
a_1^4,\ a_2^4,\ a_3^4,\ \dots,\ a_n^4,\ \dots & & A_4 \\
\dots\dots\dots\dots\dots\dots & & \dots
\end{array}$$

Перенумеруємо елементи об'єднання множин A_i в такому порядку: a_1^1 — перший, a_1^2 — другий, a_2^1 — третій, a_1^3 — четвертий, a_2^2 — п'ятий, a_3^1 — шостий, a_1^4 — сьомий і т. д. Принцип нумерації елементів множин A_i зображено за допомогою стрілок. При цьому елемент, який уже одержав свій номер раніше й повторно з'являється у процесі руху по таблиці, з подальшої нумерації виключається. У результаті кожен елемент об'єднання одержить певний номер, тобто буде задано взаємно однозначну відповідність між усіма елементами об'єднання та натуральними числами 1, 2, 3, ...

Коли принаймні одна з множин A_i зліченна чи об'єднується зліченна сукупність скінченних (попарно різних) множин, то об'єднан-

ня не може бути скінченною множиною і, отже, є зліченною множиною.

Із теореми 1.4 випливає низка цікавих наслідків.

Наслідок 1.4.1. Множина Z усіх цілих чисел зліченна.

Справді, подамо цю множину у вигляді $Z = N \cup \{0\} \cup N'$, де $N' = \{-1, -2, -3, \dots\}$ — множина від'ємних цілих чисел, яка, очевидно, зліченна.

Числова множина W називається ***щільною***, якщо для будь-якої пари чисел $a, b \in W$ $(a < b)$ існує таке число $c \in W$, що $a < c < b$.

Безпосередньо з означення випливає, що щільна множина завжди нескінченна. Більше того, для кожної пари чисел $a, b \in W$ існує безліч чисел $c \in W$, для яких виконуються нерівності $a < c < b$.

Множина Z цілих чисел, а також будь-яка її підмножина (зокрема, множина N натуральних чисел) не щільні, а множина Q раціональних чисел щільна. Для будь-яких раціональних чисел r_1 і r_2 $(r_1 < r_2)$ число $r = (r_1 + r_2)/2$ задовольняє нерівності $r_1 < r < r_2$. Зокрема, для всіх чисел r' із нескінченної множини раціональних чисел $\{r_1 + (r_2 - r_1)/2^i \mid i = 1, 2, \dots\}$ виконуються нерівності $r_1 < r' < r_2$.

Здавалося б, зі щільності множини раціональних чисел мало б випливати, що ця множина має більшу потужність, ніж множини N або Z. Однак справедливе таке твердження.

Наслідок 1.4.2. Множина Q всіх раціональних чисел зліченна.

Справді, множину Q можна подати як об'єднання таких зліченних множин:

$A_1 = \{0, 1, -1, 2, -2, 3, -3, \dots\}$ — усі цілі числа (або дроби виду $\dfrac{n}{1}$, $n \in Z$);

$$A_2 = \left\{\frac{0}{2}, \frac{1}{2}, \frac{-1}{2}, \frac{2}{2}, \frac{-2}{2}, \frac{3}{2}, \frac{-3}{2}, \dots\right\} \text{ — усі дроби виду } \frac{n}{2},\ n \in Z;$$

$$A_3 = \left\{\frac{0}{3}, \frac{1}{3}, \frac{-1}{3}, \frac{2}{3}, \frac{-2}{3}, \frac{3}{3}, \frac{-3}{3}, \dots\right\} \text{ — усі дроби виду } \frac{n}{3},\ n \in Z;$$

....

$$A_k = \left\{\frac{0}{k}, \frac{1}{k}, \frac{-1}{k}, \frac{2}{k}, \frac{-2}{k}, \frac{3}{k}, \frac{-3}{k}, \dots\right\} \text{ — усі дроби виду } \frac{n}{k},\ n \in Z,\ k \in N,$$

....

Наслідок 1.4.3. Декартів добуток $A \times B$ зліченних множин A і B — зліченна множина.

Справедливість цього твердження випливає з того, що множину всіх пар $(a, b) \in A \times B$, де $A = \{a_1, a_2, \ldots, a_n, \ldots\}$ і $B = \{b_1, b_2, \ldots, b_n, \ldots\}$ можна подати як об'єднання такої зліченної сукупності зліченних множин:

$$D_1 = \{(a_1, b_1), (a_1, b_2), \ldots, (a_1, b_n), \ldots\},$$
$$D_2 = \{(a_2, b_1), (a_2, b_2), \ldots, (a_2, b_n), \ldots\},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$D_k = \{(a_k, b_1), (a_k, b_2), \ldots, (a_k, b_n), \ldots\},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Зокрема, множина всіх точок координатної площини з раціональними координатами злічена.

Наслідок 1.4.4. Декартів добуток $P_n = A_1 \times A_2 \times \ldots \times A_n$ зліченних множин $A_1, A_2, \ldots, A_n$ є зліченною множиною для довільного n.

Доведення проведемо методом математичної індукції.

Для $n = 1$ $P_1 = A_1$, і справедливість твердження випливає з умови зліченності множини A_1. Нехай $P_{k-1} = A_1 \times A_2 \times \ldots \times A_{k-1}$ — злічена множина. Тоді зліченність множини $P_k = P_{k-1} \times A_k$ випливає з наслідку 1.4.3.

Наслідок 1.4.5. Множина P всіх многочленів від однієї змінної $p(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ з раціональними коефіцієнтами a_i, $i = 0, 1, \ldots, n$, $n = 0, 1, 2, \ldots$, злічена.

Множину P можна подати у вигляді об'єднання зліченної сукупності множин P_n, де P_n — це множина многочленів з раціональними коефіцієнтами, степінь яких не перевищує n, $n = 0, 1, 2, \ldots$. Водночас, будь-якому многочлену $p(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ з множини P_n можна поставити у відповідність кортеж $(a_0, a_1, a_2, \ldots, a_n)$, який складається з раціональних чисел a_i — коефіцієнтів цього многочлена. Ця відповідність взаємно однозначна. Отже, $P_n \sim Q^{n+1}$. Тоді з наслідків 1.4.2 й 1.4.4 випливає, що множина P_n — злічена, а тому зліченною є й множина P.

Назвемо число ***алгебричним***, якщо воно є коренем якогось многочлена з раціональними коефіцієнтами. Відомо, що кожен такий многочлен має скінченну кількість коренів (не більшу від його степеня). Тому множину всіх алгебричних чисел можна подати у ви-гляді об'єднання зліченної сукупності скінченних множин. Отже, має місце таке твердження.

Наслідок 1.4.6. Множина всіх алгебричних чисел злічена.

Наслідок 1.4.7. Множина A^* всіх слів у заданому скінченному алфавіті A злічена.

Справедливість цього твердження випливає з того, що

$$A^* = \{e\} \cup A \cup A^2 \cup A^3 \cup \ldots,$$

тобто множина A^* — це злічене об'єднання скінченних множин $\{e\}$ і A^n, де A^n — множина всіх слів довжини n в алфавіті A, $n = 1, 2, \ldots$.

1.6. Задачі і вправи

1. Довести, що після вилучення зі зліченної множини скінченної підмножини множина, що залишиться, буде зліченною.

2. Довести, що множина нескінченна тоді й тільки тоді, коли вона рівнопотужна деякій своїй власній підмножині.

3. Довести, що непорожня множина A злічена або скінченна тоді й тільки тоді, коли вона є множиною значень якогось відображення з N в A.

4. Довести, що множина всіх скінченних послідовностей, що складаються з елементів якоїсь зліченної множини, злічена.

5. Довести, що множина всіх скінченних підмножин зліченної множини злічена.

6. Довести, що будь-яка множина відкритих інтервалів на дійсній прямій, які попарно не перетинаються, скінченна або злічена.

7. Довести, що множина точок розриву монотонної функції на дійсній прямій є скінченною або зліченною.

1.8. Незліченні множини

Із висловленого вище припущення про рівнопотужність усіх нескінченних множин логічно випливають такі питання: чи всі нескінченні множини зліченні, або чи існують нескінченні множини, які не будуть зліченними? Факт існування нескінченних множин, які не є зліченними (***незліченних*** множин), вперше встановив Г. Кантор за допомогою запропонованого ним *діагонального методу*, який набув згодом фундаментального значення в різних розділах математики. Зокрема, цей метод лежить в основі доведення наступної важливої теореми, що належить Г. Кантору.

Теорема 1.5. Множина всіх дійсних чисел з інтервалу (0, 1) незлічена.

Доведення. Відомо, що кожному дійсному числу з інтервалу (0, 1) можна поставити у відповідність нескінченний десятковий дріб

$0,a_1 a_2 a_3\ldots$. Для ірраціональних чисел такий дріб неперіодичний. Для кожного раціонального числа, зображуваного скінченним десятковим дробом, із двох можливих варіантів запису його у вигляді нескінченного періодичного десяткового дробу (з періодом 0 або 9) зафіксуємо період 9. Наприклад, число 0,123 (або 0,123000…) будемо записувати у вигляді 0,122999…, а число 0,7 — у вигляді 0,699… . Запропонована відповідність взаємно однозначна.

Доведемо теорему методом від супротивного.

Припустімо, що сформульоване твердження хибне, і множина всіх дійсних чисел з інтервалу (0, 1) зліченна. Отже, існує нумерація цих чисел $x_1, x_2, \ldots, x_n, \ldots$. Перепишемо всі числа з інтервалу (0, 1) у вигляді нескінченних десяткових дробів у порядку їх нумерації:

$$
\begin{array}{l}
x_1 = 0,\ a_{11}\ a_{12}\ a_{13} \ldots a_{1n} \ldots, \\
\qquad\qquad \searrow \\
x_2 = 0,\ a_{21}\ a_{22}\ a_{23} \ldots a_{2n} \ldots, \\
\qquad\qquad\quad \searrow \\
x_3 = 0,\ a_{31}\ a_{32}\ a_{33} \ldots a_{3n} \ldots, \\
\qquad\qquad\qquad \searrow \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_n = 0,\ a_{n1}\ a_{n2}\ a_{n3} \ldots a_{nn} \ldots, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{array}
$$

Рухаючись по діагоналі (вказаній стрілками), утворимо новий нескінченний десятковий дріб $0,b_1 b_2 \ldots b_n \ldots$ такий, що $b_1 \neq a_{11}$, $b_2 \neq a_{22}$, … …, $b_n \neq a_{nn}$, … . Щоб уникнути ситуації подання того самого раціонального числа у двох формах, оберемо значення цифр b_i так, щоб $b_i \neq 0$ і $b_i \neq 9$, $i = 1, 2, \ldots$. Утворений дріб є записом якогось дійсного числа y з інтервалу (0, 1), однак він не належить розглядуваній зліченній множині. Справді, побудований дріб відрізняється від кожного з дробів нашої нумерації $x_1, x_2, \ldots, x_n, \ldots$ принаймні однією цифрою. Наприклад, $y \neq x_n$ тому, що дроби y та x_n відрізняються принаймні n-ю цифрою після коми ($n = 1, 2, \ldots$). З одержаної суперечності випливає, що не існує переліку для множини всіх дійсних чисел з інтервалу (0, 1). Отже, припущення щодо її зліченності хибне, і розглядувана множина незліченна. Теорему доведено.

Будь-яка множина, рівнопотужна множині всіх дійсних чисел з інтервалу (0, 1), називається ***континуальною***, або множиною ***потужності континуум***.

Із наведених вище прикладів 1.12 (3, 4) та зауваження про рівнопотужність усіх інтервалів і відрізків дійсної прямої, а також твердження про рівнопотужність будь-якого відрізка та всієї дійсної прямої випливає, що всі ці множини точок континуальні.

Теорема 1.6. Якщо M — незліченна множина, а A — скінченна чи зліченна підмножина множини M, то множини $M \setminus A$ і M рівнопотужні, тобто $M \setminus A \sim M$.

Доведення. Множина $M \setminus A$ незліченна. Якби множина $M' = M \setminus A$ була зліченною, то за теоремою 1.4 множина $M = M' \cup A$ також була б зліченною, що суперечило б умові теореми. За теоремою 1.2 множина M' містить зліченну підмножину B ($B \subseteq M \setminus A$). По-значимо $C = (M \setminus A) \setminus B$, тоді маємо $M \setminus A = B \cup C$ і $M = (A \cup B) \cup C$. Множина $A \cup B$ зліченна. Тоді з рівнопотужностей $B \sim (A \cup B)$ та $C \sim C$, а також того, що $C \cap B = \varnothing$ і $C \cap (A \cup B) = \varnothing$, випливає співвідношення $B \cup C \sim (A \cup B) \cup C$, тобто $M \setminus A \sim M$.

Сформулюємо кілька наслідків, що випливають з доведених теорем.

Наслідок 1.6.1. Якщо M — нескінченна множина, а множина A скінченна чи зліченна, то $M \cup A \sim M$.

Будемо вважати, що $M \cap A = \varnothing$. Якщо $M \cap A \neq \varnothing$, то в доведенні можна використати таку скінченну чи зліченну множину $A' = A \setminus M$, що $M \cup A = M \cup A'$ і $M \cap A' = \varnothing$.

Якщо множина M зліченна, то $M \cup A$ — також зліченна множина (теорема 1.4), отже, $M \cup A \sim M$.

Якщо множина M незліченна, то $M \cup A$ — також незліченна. Тоді за теоремою 1.6 $(M \cup A) \setminus A \sim M \cup A$, тобто $M \sim M \cup A$.

Наслідок 1.6.2. Множина всіх ірраціональних чисел континуальна.

Число, яке не є коренем жодного многочлена з раціональними коефіцієнтами, називається ***трансцендентним***.

Наслідок 1.6.3. Множина всіх трансцендентних чисел континуальна.

Справедливість наслідків 1.6.2 та 1.6.3 випливає з континуальності множин R і C відповідно всіх дійсних і комплексних чисел, зліченності множин усіх раціональних і всіх алгебричних чисел, а також теореми 1.6.

Із доведених теорем випливає, зокрема, рівнопотужність інтервалів (0, 1), [0, 1), (0, 1] і [0, 1].

Сформульована нижче теорема встановлює певний зв'язок між зліченними та континуальними множинами. Для її доведення теж застосовано діагональний метод Кантора.

Теорема 1.7. Множина $\beta(A)$ всіх підмножин зліченної множини A має потужність континуум.

Доведення. Оскільки всі зліченні множини рівнопотужні множині N натуральних чисел, то достатньо довести континуальність булеана $\beta(N)$ множини N. Маючи взаємно однозначну відповідність між множиною N і якоюсь множиною A, неважко побудувати взаємно однозначну відповідність між їх булеанами $\beta(N)$ і $\beta(A)$.

Доведемо теорему методом від супротивного. Припустімо, що множина $\beta(N)$ злічення й існує нумерація всіх її елементів, тобто $\beta(N) = \{M_1, M_2, \ldots, M_k, \ldots\}$, де $M_k \subseteq N$, $k = 1, 2, \ldots$. Поставимо у відповідність кожній множині M_k послідовність t_k з нулів і одиниць $m_1^{(k)}, m_2^{(k)}, \ldots, m_i^{(k)}, \ldots$ за таким законом: $m_i^{(k)} = 1$, якщо $i \in M_k$, і $m_i^{(k)} = 0$, якщо $i \notin M_k$. Ця відповідність взаємно однозначна.

Розмістимо всі елементи множини $\beta(N)$ і відповідні їм послідовності згідно з нумерацією:

$$\begin{array}{l}
M_1 \;—\; m_1^{(1)}, m_2^{(1)}, \ldots, m_k^{(1)}, \ldots \\
M_2 \;—\; m_1^{(2)}, m_2^{(2)}, \ldots, m_k^{(2)}, \ldots \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
M_k \;—\; m_1^{(k)}, m_2^{(k)}, \ldots, m_k^{(k)}, \ldots \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{array}$$

За допомогою діагонального методу Кантора побудуємо таку нову послідовність L з нулів і одиниць $l_1, l_2, \ldots, l_k, \ldots$, що $l_k \neq m_k^{(k)}$, тобто $l_k = 1$, якщо $m_k^{(k)} = 0$, і $l_k = 0$, якщо $m_k^{(k)} = 1$, $k = 1, 2, 3, \ldots$.

Послідовності L відповідає якась підмножина $M \subseteq N$, а саме $M = \{n \mid l_n = 1, n = 1, 2, \ldots\}$. Підмножина M не входить до зазначеного переліку $M_1, M_2, \ldots, M_k, \ldots$, бо послідовність L відрізняється від кожної з послідовностей t_k принаймні в одній k-й позиції. Отже, і множина M відрізняється від кожної з множин M_k, $k = 1, 2, \ldots$. Ця суперечність означає, що не існує переліку для елементів множини $\beta(N)$. Тому множина $\beta(N)$ незліченна.

Крім того, кожній послідовності t_k можна поставити у відповідність нескінченний двійковий дріб $0{,}m_1^{(k)}m_2^{(k)}\ldots m_k^{(k)}\ldots$, який зображує якесь дійсне число з інтервалу $[0, 1]$ у двійковій системі числення. І навпаки,

будь-яке число з інтервалу $[0, 1]$ можна однозначно записати у вигляді нескінченного двійкового дробу.

Виняток становлять числа зі зліченної множини раціональних чисел, які записуються за допомогою скінченних двійкових дробів, внаслідок чого вони можуть мати дві різні форми зображення у вигляді нескінченних двійкових дробів — з періодами 0 або 1. Кожному з таких чисел відповідають дві різні послідовності t' і t'', а отже, і два різні елементи множини $\beta(N)$: один — для зображення з періодом 0, другий — з періодом 1. Позначимо через T множину підмножин множини N, зіставлюваних за допомогою побудованої вище відповідності нескінченним двійковим дробам з періодом 0, $T \subseteq \beta(N)$. Тоді існує взаємно однозначна відповідність між множиною всіх дійсних чисел з інтервалу $[0, 1]$ і множиною $\beta(N) \setminus T$. Однак, оскільки множина T зліченна, то за теоремою 1.6 маємо $\beta(N) \sim \beta(N) \setminus T$. Отже, множина $\beta(N)$, а значить, і множина $\beta(A)$ для будь-якої зліченної множини A мають потужність континуум.

1.7. Задачі і вправи

1. Встановити взаємно однозначну відповідність між точками
(а) двох квадратів;
(б) квадрата $(a, b) \times (a, b)$ та площини;
(в) сфери, з якої вилучено одну точку, і площиною;
(г) прямої та півпрямої (променя);
(д) кола без однієї точки і прямої.

2. Довести, що об'єднання скінченної чи зліченної сукупності континуальних множин є континуальною множиною.

3. Довести, що множина всіх зліченних послідовностей натуральних чисел має потужність континуум.

4. Знайти помилку в наведених міркуваннях. Позначимо через A_k множину всіх k-елементних підмножин множини натуральних чисел N. Ця множина зліченна. Отже, зліченним є й булеан $\beta(N)$, тому що $\beta(N) = A_0 \cup A_1 \cup A_2 \cup \ldots \cup A_k \cup \ldots$.

5. Чи можна накреслити на площині континуальну множину кіл, жодні два з яких не перетинаються?

6. Чи можна накреслити на площині континуальну множину кругів, жодні два з яких не перетинаються?

7. Чи можна накреслити на площині континуальну множину літер А, жодні дві з яких не мають спільних точок?

8. Чи можна накреслити на площині континуальну множину літер Г, жодні дві з яких не перетинаються?

1.9. Кардинальні числа

Нехай A — якась множина і $S = \{B \mid B \sim A\}$ — сукупність усіх множин, рівнопотужних множині A. Всі множини з S рівнопотужні. ***Кардинальним числом*** (позначається $|A|$, $\overline{\overline{A}}$ або Card (A)) називають певний об'єкт для позначення потужності будь-якої множини із сукупності S.

Зокрема, для скінченної множини A кардинальним числом $|A|$ є натуральне число, що позначає кількість елементів будь-якої з множин сукупності S. Отже, можна вважати, що кардинальне число є узагальненим поняттям числа елементів.

Природно виникає питання про порівняння кардинальних чисел нескінченних множин.

Нехай A і B — нескінченні множини. Тоді логічно можливі такі чотири випадки:

1) існує взаємно однозначна відповідність між A і B, тобто $A \sim B$ і $|A| = |B|$;

2) існує взаємно однозначна відповідність між множиною A та якоюсь підмножиною B' множини B. Тоді кажуть, що потужність множини A ***не більша*** від потужності множини B, і записують $|A| \le |B|$;

3) множина A рівнопотужна якійсь підмножині множини B і, навпаки, множина B рівнопотужна якійсь підмножині множини A, тобто $A \sim B' \subseteq B$ та $B \sim A' \subseteq A$. За теоремою Кантора—Бернштейна, доведення якої наведено нижче, у цьому разі $A \sim B$, тобто $|A| = |B|$.

4) не існує взаємно однозначної відповідності між множиною A і жодною підмножиною множини B, а також між множиною B і жодною підмножиною множини A. Із цієї ситуації випливало б, що потужності множин A і B непорівнювані між собою.

Однак глибші дослідження в теорії множин показали, що, спираючись на *аксіому вибору* (див. розд. 1.12), можна довести неможливість четвертого випадку.

Отже, потужності будь-яких двох множин A та B завжди порівнювані між собою. Тому для кардинальних чисел $|A|$ і $|B|$ довільних множин A та B виконується принаймні одне з трьох співвідношень: $|A| = |B|$, $|A| \le |B|$ або $|B| \le |A|$. Наприклад, неважко переконатися, що коли $A \subseteq B$, то $|A| \le |B|$.

Якщо $|A| \le |B|$ та множина A нерівнопотужна множині B, то писатимемо $|A| < |B|$.

Теорема 1.8 (***Кантора—Бернштейна***). Якщо множина A рівнопотужна якійсь підмножині B_1 множини B, $A \sim B_1 \subseteq B$ та водночас множина B рівнопотужна якійсь підмножині A_1 множини A, $B \sim A_1 \subseteq A$, то множини A і B рівнопотужні.

Доведення. Позначимо через f бієкцію між множинами A і B_1, а через g — бієкцію між B й A_1. Нехай x — довільний елемент множини A. Розглянемо послідовність $x_1, x_2, \ldots$, утворену за таким правилом. Покладемо $x_1 = x$. Якщо елемент $x_{2k-1} \in A$ визначено, то вважатимемо, що $x_{2k} = g^{-1}(x_{2k-1})$ (у тому разі, коли такий елемент існує, тобто якщо $g^{-1}(x_{2k-1}) \ne \varnothing$ чи $x_{2k-1} \in A_1$); якщо елемент $x_{2k} \in B$ визначено, то покладемо $x_{2l+1} = f^{-1}(x_{2l})$ (у разі, коли такий елемент існує, тобто якщо $f^{-1}(x_{2k}) \ne \varnothing$ чи $x_{2k} \in B_1$), $k \in N$.

Якщо для якогось n елемента x_{n+1} не існує, то будована за x послідовність скінченна, і число n називатимемо *порядком* елемента x. Якщо ж ця послідовність нескінченна, то називатимемо x *елементом нескінченного порядку*.

Таким чином множина A розбивається на три підмножини: $A^{(1)}$ — елементів непарного порядку, $A^{(2)}$ — елементів парного порядку й $A^{(0)}$ — елементів нескінченного порядку. За допомогою аналогічної процедури множину B можна розбити на відповідні підмножини $B^{(1)}$, $B^{(2)}$ та $B^{(0)}$. Неважко переконатись, що $f(A^{(1)}) = B^{(2)}$, $g(B^{(1)}) = A^{(2)}$ (або $g^{-1}(A^{(2)}) = B^{(1)}$) й $f(A^{(0)}) = B^{(0)}$.

Відповідність $h = \left(f \cap \left((A^{(1)} \cup A^{(0)}) \times B^{(1)}\right)\right) \cup \left(g^{-1} \cap (A^{(2)} \times B)\right)$, що збігається з f для елементів з $A^{(1)} \cup A^{(0)}$ та збігається з g^{-1} для $A^{(2)}$, буде бієкцією між множинами A і B. Теорему доведено.

Наслідок 1.8.1. Якщо виконуються включення $A_2 \subseteq A_1 \subseteq A$ і $A_2 \sim A$ $(|A_2| = |A|)$, то $A_1 \sim A$ $(|A_1| = |A|)$.

Твердження випливає з того, що $A \sim A_2 \subseteq A_1$ і $A_1 \sim A_1 \subseteq A$.

Для кардинальних чисел зліченних і континуальних множин, враховуючи їх поширеність та популярність в сучасній математиці, введено спеціальні позначення. Так, кардинальне число множини N натуральних чисел, а значить, і будь-якої зліченної множини, позначають $\aleph_0$ (читається "алеф-нуль"). Кардинальне число континуальних множин позначають через **c** або $\aleph_1$ ("алеф-один").

Якщо порівняти доведення теорем 1.1 і 1.7, то неважко помітити аналогію у встановленні взаємно однозначної відповідності між підмножинами множини A та двійковими послідовностями (скінченними в теоремі 1.1 і нескінченними в теоремі 1.7). Враховуючи цю аналогію, часто записують співвідношення $|\beta(A)| = 2^{|A|}$ як для скінченних, так і для нескінченних множин. Деякі автори для булеана множини A використовують позначення 2^A, що дає змогу записувати останню рівність у вигляді $|2^A| = 2^{|A|}$. Зокрема, за теоремою 1.7 $\aleph_1 = 2^{\aleph_0}$ [1].

Наступне питання, що виникло в теорії множин: чи існує найбільше кардинальне число, тобто чи існує множина найбільшої потужності? Негативну відповідь на це питання дає наступна важлива теорема, доведення якої належить Г. Кантору.

Теорема 1.9. Потужність множини $\beta(A)$ всіх підмножин будь-якої непорожньої множини A більша від потужності множини A: $|A| < |\beta(A)|$.

Доведення. Оскільки існує тривіальна взаємно однозначна відповідність f між множиною A та підмножиною множини $\beta(A)$: $f = \{(a,\{a\}) \mid a \in A,\ \{a\} \in \beta(A)\}$, то $|A| \le |\beta(A)|$, і потрібно довести, що множини A та $\beta(A)$ нерівнопотужні.

Доведення проведемо методом від супротивного. Припустімо, що існує взаємно однозначна відповідність g між множинами A і $\beta(A)$: $g = \{(b, B) \mid b \in A$ і $B \in \beta(A)\}$. У кожній парі відповідності g перша координата b — це елемент множини A, а друга координата B — відповідна елементу b підмножина множини A. Тому для кожної пари $(b, B) \in g$ виконується одне з двох співвідношень: або $b \in B$, або $b \notin B$. Побудуємо нову множину $K = \{b \mid b \in A$ і $b \notin B$ для $(b, B) \in g\}$. З того, що $\varnothing \in \beta(A)$, випливає, що $K \ne \varnothing$.

Оскільки K є підмножиною множини A $(K \in \beta(A))$, то при взаємно однозначній відповідності g підмножина K відповідає якомусь елементу $k \in A$, тобто існує пара $(k, K) \in g$. Тоді для елемента $k \in A$ та підмножини $K \subseteq A$ можливі дві ситуації: або $k \in K$, або $k \notin K$.

Нехай $k \in K$, тоді з умови $(k, K) \in g$ і правила побудови множини K випливає, що $k \notin K$. З іншого боку, якщо припустити, що $k \notin K$, то з $(k, K) \in g$ і правила побудови множини K випливає $k \in K$.

Одержана суперечність доводить неможливість встановлення взаємно однозначної відповідності між A і $\beta(A)$. Отже, $|A| < |\beta(A)|$.

Наслідок 1.9.1. Не існує множини, яка має найбільшу потужність, тобто не існує найбільшого кардинального числа.

Справді, розглянувши множини N, $\beta(N)$, $\beta(\beta(N))$, ..., одержимо послідовність відповідних кардинальних чисел $\aleph_0$, $\aleph_1 = 2^{\aleph_0}$, $\aleph_2 = 2^{\aleph_1}$, ..., яка нескінченно зростає.

На закінчення зупинімося ще на одній цікавій класичній проблемі теорії множин, сформульованій ще в 1884 р. Г. Кантором, — ***гіпотезі континуума***, згідно з якою не існує множини, кардинальне число $\aleph$ якої лежить між $\aleph_0$ і $\aleph_1$, тобто $\aleph_0 < \aleph < \aleph_1$.

Цю гіпотезу можна узагальнити до *узагальненої гіпотези континуума*: для довільного кардинального числа $\aleph$ якоїсь нескінченної множини з нерівності $\aleph' > \aleph$ для будь-якого кардинального числа $\aleph'$ випливає $\aleph' \ge 2^{\aleph}$.

Проблему гіпотези континуума майже вісім десятків років намагалися розв'язати найкращі математики світу. І лише в 1963 р. тридцятирічний американський математик Пол Коен довів, що гіпотезу континуума не можна ні довести, ні спростувати, виходячи з аксіом теорії множин. Отже, прийняття або відхилення гіпотези континуума однаково законні, що дає змогу будувати дві різні несуперечливі теорії множин.

1.8. Задачі і вправи

1. Довести, що зі співвідношень $|A| \le |B|$ та $|B| \le |A|$ випливає $|A| = |B|$.

2. Довести, що множини точок квадрата $(a, b) \times (a, b)$ та відрізка (c, d) рівнопотужні.

3. Довести, що $R^n \sim R^m$ $(n, m \ge 1)$.

4. Нехай усі множини A_i, $i = 1, 2, \ldots, n$ континуальні. Довести, що множина $A_1 \times A_2 \times \ldots \times A_n$ також континуальна.

5. Довести таке твердження:
(а) якщо $A \subseteq B$, то $|A| \le |B|$;
(б) якщо $|A| \le |B|$ і $|B| \le |C|$, то $|A| \le |C|$;
(в) якщо $A_1 \sim A_2$, $B_1 \sim B_2$ і $|A_1| \le |B_1|$, то $|A_2| \le |B_2|$;
(г) якщо існує відображення з A на B, то $|B| \le |A|$.

6. Довести, що для довільної множини A виконується нерівність $|A| \le |A \times A|$.

7. Нехай $|A_1| \le |B_1|$ і $|A_2| \le |B_2|$. Довести таке твердження:
(а) якщо $B_1 \cap B_2 = \varnothing$, то $|A_1 \cup A_2| \le |B_1 \cup B_2|$;
(б) $|A_1 \times A_2| \le |B_1 \times B_2|$.

8. Нехай A — довільна непорожня множина, а f — довільне відображення множини A в $\beta(A)$. Довести, що:
(а) $f(A) \ne \beta(A)$; (б) $|f(A)| < |\beta(A)|$.

9. Якою є потужність множини:
(а) усіх відображень типу $N \to N$;
(б) усіх монотонних відображень типу $N \to N$;

(в) усіх нескінченних послідовностей дійсних чисел;

(г) усіх функцій типу $N \to \{0, 1\}$;

(д) усіх відображень типу $N \to \{0, 1\}$;

(е) усіх неперервних відображень типу $R \to R$;

(є) усіх монотонних відображень типу $R \to R$;

(ж) усіх відображень типу $N \to R$;

(з) усіх відображень (функцій) типу $R \to N$?

10. Довести, що множина всіх дійсних відображень (функцій), заданих на інтервалі (0, 1), має більшу потужність, ніж **c**.

1.10. Відношення. Властивості відношень

Підмножина R декартового степеня M^n множини M називається ***n-місним*** (***n-арним***) ***відношенням*** на множині M. Кажуть, що елементи $a_1, a_2, \ldots, a_n \in M$ перебувають у відношенні R, якщо $(a_1, a_2, \ldots, a_n) \in R$.

У разі $n = 1$ відношення $R \subseteq M$ називають ***одномісним***, або ***унарним***. Таке відношення часто називають також *ознакою*, або *характеристичною властивістю* елементів множини M. Кажуть, що елемент $a \in M$ має ознаку R, якщо $a \in R$ і $R \subseteq M$. Наприклад, ознаки "непарність" і "кратність 7" виділяють із множини N натуральних чисел відповідно унарні відношення $R' = \{2k-1 \mid k \in N\}$ і $R'' = \{7k \mid k \in N\}$.

Найпопулярнішими в математиці є ***двомісні*** (***бінарні***) відношення, на вивченні властивостей яких зупинимося детальніше. Далі скрізь під словом "відношення" розумітимемо бінарне відношення. Якщо елементи $a, b \in M$ перебувають у відношенні R, тобто $(a, b) \in R$, то це часто записують також у вигляді $a\,R\,b$. Зауважимо, що бінарні відношення іноді розглядають як окремий випадок відповідностей (а саме — як відповідності між однаковими множинами), тому багато означень і понять для відношень подібні до аналогічних означень і понять для відповідностей.

Приклад 1.13. Наведемо приклади бінарних відношень на різних множинах.

1. Відношення на множині N натуральних чисел:

R_1 — відношення "менше або дорівнює", тоді $4\,R_1\,9$, $5\,R_1\,5$ і $1\,R_1\,m$ для будь-якого $m \in N$;

R_2 — відношення "ділиться на", тоді $24\,R_2\,3$, $49\,R_2\,7$ і $m\,R_2\,1$ для будь-якого $m \in N$;

R_3 — відношення "є взаємно простими", тоді $15\,R_3\,8$, $366\,R_3\,121$, $1001\,R_3\,612$;

R_4 — відношення "складаються з однакових цифр", тоді $127\,R_4\,721$, $230\,R_4\,302$, $3231\,R_4\,3213311$.

2. Відношення на множині точок координатної площини R^2:

R_5 — відношення "лежать на однаковій відстані від початку координат", тоді $(3, 2)\ R_5\,(\sqrt{5}, -\sqrt{8})$, $(0, 0)\ R_5\,(0, 0)$;

R_6 — відношення "симетричні відносно осі ординат", тоді $(-3, 2)\,R_6\,(3, 2)$, $(1, 7)\,R_6\,(-1, 7)$ і взагалі $(a, b)\,R_6\,(-a, b)$ для будь-яких $a, b \in R$;

R_7 — відношення "менше або дорівнює". Вважаємо, що $(a, b)\ R_7\,(c, d)$, якщо $a \le c$ і $b \le d$. Зокрема, $(1, 7)\ R_7\,(20, 14)$, $(-12, 4)\ R_7\,(0, 17)$, а пари $(2, -7)$ і $(1, 4)$ та $(3, -5)$ і $(-3, 2)$ не належать відношенню R_7.

3. Відношення на множині студентів вищого навчального закладу:

R_8 — відношення "є однокурсником",

R_9 — відношення "молодший за віком від".

Відношення можна задавати тими самими способами, що й звичайні множини. Наприклад, якщо множина M скінченна, то довільне відношення R на M можна задати списком пар елементів, що перебувають у відношенні R.

Крім того, зручно задавати бінарне відношення R на скінченній множині $M = \{a_1, a_2, \ldots, a_n\}$ за допомогою так званої ***матриці бінарного відношення***. Це квадратна матриця C порядку n, у якій елемент c_{ij}, що стоїть на перетині i-го рядка та j-го стовпчика, означають так: $c_{ij} = 1$, якщо $a_i\,R\,a_j$, і $c_{ij} = 0$, в іншому разі.

Відношення можна задавати також за допомогою графіків і діаграм. ***Графік відношення*** означають і будують так само, як і графік відповідності. Поняття діаграми (або графа) відношення також можна означити аналогічно до відповідності. Однак частіше ***діаграму*** (***граф***) ***відношення*** R на скінченній множині $M = \{a_1, a_2, \ldots, a_n\}$ означають так. Поставимо у взаємно однозначну відповідність елементам множини M якісь точки площини. Із точки a_i до точки a_j проводимо напрямлену лінію (стрілку) у вигляді відрізка чи кривої тоді й тільки тоді, коли $a_i\,R\,a_j$. Зокрема, якщо $a_i\,R\,a_i$, то відповідна стрілка, що веде з a_i в a_i, називається ***петлею***.

Приклад 1.14. Для множини $M = \{2, 7, 36, 63, 180\}$ матриці відношень R_1, R_2, R_3 з останнього прикладу мають такий вигляд:

$$C_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Діаграми (графи) відношень R_1, R_2, R_3 подано на рис. 1.4.

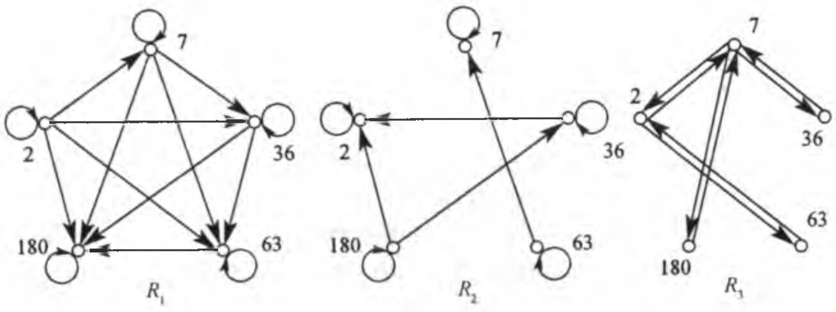


Рис. 1.4

Оскільки відношення на M є підмножинами множини M^2, то для них означені всі відомі теоретико-множинні операції. Наприклад, перетином відношень "більше або дорівнює" і "менше або дорівнює" є відношення "дорівнює", об'єднанням відношень "менше" і "більше" є відношення "не дорівнює", доповненням відношення "ділиться на" є відношення "не ділиться на" тощо. Стосовно операції доповнення для відношень на множині M універсальною множиною є M^2.

Аналогічно відповідностям для відношень можна означити поняття оберненого відношення та композиції відношень.

Відношення R^{-1} називається ***оберненим*** до відношення R, якщо $b\,R^{-1}\,a$ тоді й тільки тоді, коли $a\,R\,b$. Очевидно, що $\left(R^{-1}\right)^{-1} = R$.

Наприклад, для відношення "більше або дорівнює" оберненим відношенням є "менше або дорівнює", для відношення "ділиться на" — відношення "є дільником".

Композицією відношень R_1 і R_2 на множині M (позначається $R_1 \circ R_2$) називається таке відношення R на M, що $a\,R\,b$ тоді й тільки тоді, коли існує елемент $c \in M$, для якого виконується $a\,R_1\,c$ і $c\,R_2\,b$.

Наприклад, композицією відношень R_1 — "є сином" і R_2 — "є братом" на множині чоловіків є відношення $R_1 \circ R_2$ — "є небожем".

Для відношення R на множині M через $R^{(k)}$ позначимо відношення $R \circ R \circ \ \ldots\ \circ R$ (k разів). Вважаємо, що $R^{(0)} = i_M$ і $R^{(1)} = R$.

Наведемо *властивості*, за якими класифікують відношення.

Нехай R — відношення на множині M.

1. Відношення R називається ***рефлексивним***, якщо для всіх $a \in M$ виконується $a\,R\,a$.

Відношення R_1, R_2, R_4, R_5, R_7 — рефлексивні.

2. Відношення R називають ***антирефлексивним*** (***іррефлексивним***), якщо для жодного $a \in M$ не виконується $a\,R\,a$.

Відношення "більше", "менше", "є сином" антирефлексивні, а відношення R_6 ні рефлексивне, ні антирефлексивне.

Усі елементи головної діагоналі матриці C для рефлексивного відношення на скінченній множині M дорівнюють 1, а для антирефлексивного — 0.

3. Відношення R називається ***симетричним***, якщо для всіх $a, b \in M$ таких, що $a\,R\,b$, маємо $b\,R\,a$.

4. Відношення R називається ***антисиметричним***, якщо для всіх $a, b \in M$ таких, що $a\,R\,b$ і $b\,R\,a$, маємо $a = b$.

Наприклад, відношення R_3, R_4, R_5, R_6, R_8 — симетричні, а відношення R_1, R_2, R_7 — антисиметричні.

Неважко переконатися, що відношення R симетричне тоді й тільки тоді, коли $R = R^{-1}$.

5. Відношення R називається ***транзитивним***, якщо зі співвідношень $a\,R\,b$ та $b\,R\,c$ випливає $a\,R\,c$.

Наприклад, відношення R_1, R_2, R_4, R_5, R_7, R_8, R_9 — транзитивні, а відношення R_3, R_6 — не транзитивні.

Неважко переконатися, що відношення R транзитивне тоді й тільки тоді, коли $R \circ R \subseteq R$.

Якщо відношення R має будь-яку з наведених вище властивостей, то обернене відношення R^{-1} також має ту саму властивість. Отже, операція обернення зберігає всі п'ять властивостей відношень.

Для довільного відношення R означимо нову операцію. Відношення R^+ називається ***транзитивним замиканням*** відношення R на M, якщо $a\,R^+\,b$, $a, b \in M$, тоді й тільки тоді, коли у множині M існує така послідовність елементів $a_1, a_2, \ldots, a_k$, що $a_1 = a$, $a_k = b$ і $a_1 R a_2$, $a_2 R a_3$, … …, $a_{k-1} R a_k$. Зокрема, k може дорівнювати 2; отже, якщо $a R b$, то $a R^+ b$. Тому $R \subseteq R^+$.

Наприклад, нехай M — множина точок на площині й aRb, $a, b \in M$, якщо точки a і b з'єднано відрізком. Тоді $c R^+ d$, коли існує ламана лінія, що з'єднує точки c і d, $c, d \in M$.

Можна довести, що відношення R транзитивне тоді й тільки тоді, коли $R^+ = R$. Крім того, справджується рівність

$$R^+ = R^{(1)} \cup R^{(2)} \cup \ldots \cup R^{(k)} \cup \ldots .$$

Відношення $R^+ \cup i_M$ позначають R^* і називають ***рефлексивним транзитивним замиканням*** відношення R на множині M.

Відношення R на множині M називається ***толерантним*** (***відношенням толерантності***, або просто ***толерантністю***), якщо воно рефлексивне і симетричне.

Деякі відношення займають особливе місце в математиці. Розглянемо їх окремо в наступних розділах.

1.9. Задачі і вправи

1. Нехай на множині всіх людей P означено такі відношення:
$F = \{(x, y) \mid x, y \in P$ і x — батько $y\}$ і $D = \{(x, y) \mid x, y \in P$ і x — донька $y\}$. Описати таке відношення:
(а) $F \circ F$; (в) $F \circ D$; (д) $D \circ F^{-1}$; (є) $F^{-1} \circ D^{-1}$;
(б) $D \circ D$; (г) $D \circ F$; (е) $F^{-1} \circ D$; (ж) $D^{-1} \circ F$.

2. Довести, що для довільного відношення R на множині M виконується
(а) $\mathrm{Pr}_2 R = \mathrm{Pr}_1 R^{-1}$; (в) $\mathrm{Pr}_1 R = M \Leftrightarrow i_M \subseteq R \circ R^{-1}$;
(б) $\mathrm{Pr}_1 R = \mathrm{Pr}_2 R^{-1}$; (г) $\mathrm{Pr}_2 R = M \Leftrightarrow i_M \subseteq R^{-1} \circ R$.

3. Довести, що для довільних відношень R_1 і R_2 виконується
(а) $(R_1 \circ R_2)^{-1} = R_2^{-1} \circ R_1^{-1}$;
(б) $R_1 \circ R_2 = \varnothing \Leftrightarrow \mathrm{Pr}_2 R_1 \cap \mathrm{Pr}_1 R_2 = \varnothing$.

4. Для яких відношень виконується рівність $R^{-1} = \mathrm{R}$?

5. На множині $M = \{1, 2, 3, 4\}$ задано відношення
$R_1 = \{(1, 1), (1, 3), (2, 2), (2, 4), (3, 1), (3, 3), (4, 2), (4, 4)\}$;
$R_2 = \{(1, 1), (2, 2), (2, 3), (2, 4), (3, 1), (3, 3), (4, 1), (4, 4)\}$;
$R_3 = \{(1, 3), (1, 4), (2, 1), (1, 2), (3, 1), (3, 3), (4, 1)\}$;
$R_4 = \{(1, 1), (1, 2), (1, 4), (2, 2), (2, 4), (3, 3), (3, 4), (4, 2), (4, 3), (4, 4)\}$;
$R_5 = \{(1, 2), (1, 3), (2, 3), (2, 4), (4, 1), (4, 3)\}$.
Визначити, які з цих відношень
(а) рефлексивні; (в) симетричні; (д) транзитивні;
(б) антирефлексивні; (г) антисиметричні; (е) толерантні.
Побудувати графіки, графи і матриці заданих відношень.

6. Проінтерпретуйте властивості відношень за допомогою їх матриць, графіків і діаграм.

7. Довести, що коли $R_1 \subseteq R_2$, то для довільного відношення Q виконується
(а) $Q \circ R_1 \subseteq Q \circ R_2$; (б) $R_1 \circ Q \subseteq R_2 \circ Q$; (в) $R_1^{-1} \subseteq R_2^{-1}$.

8. Довести, що відношення R на множині M рефлексивне тоді й лише тоді, коли
(а) $i_M \subseteq R$; (б) $i_M \cap R = i_M$; (в) $i_M \cup R = R$.

9. Навести приклад двох антирефлексивних відношень R_1 і R_2 на множині $M = \{1, 2, 3\}$, композиція $R_1 \circ R_2$ яких буде рефлексивним відношенням.

10. Довести, що відношення R симетричне тоді й тільки тоді, коли
(а) $R = R^{-1}$; (б) $R^{-1} \subseteq R$; (в) $R \subseteq R^{-1}$.

11. Довести, що композиція $R_1 \circ R_2$ симетричних відношень R_1 і R_2 є симетричним відношенням тоді й тільки тоді, коли $R_1 \circ R_2 = R_2 \circ R_1$.

12. Довести, що відношення R на множині M антисиметричне тоді й тільки тоді, коли $R \cap R^{-1} \subseteq i_M$.

13. Довести, що об'єднання $R_1 \cup R_2$ антисиметричних відношень R_1 і R_2 на множині M є антисиметричним відношенням тоді й лише тоді, коли $R_1 \cap R_2^{-1} \subseteq i_M$.

14. Довести, що відношення R на множині M транзитивне тоді й тільки тоді, коли $R \circ R \subseteq R$.

15. Довести, що композиція $R_1 \circ R_2$ транзитивних відношень R_1 і R_2 є транзитивним відношенням, якщо $R_1 \circ R_2 = R_2 \circ R_1$.

16. Знайти помилку в наведених міркуваннях. Якщо R — симетричне і транзитивне відношення на множині M, то воно рефлексивне, бо з того, що $(a, b) \in R$, послідовно випливає $(b, a) \in R$ і $(a, a) \in R$.

17. Довести, що відношення $T = \{(x, y) \mid |x - y| < 1, x, y \in R\}$ на множині дійсних чисел R є толерантним.

18. Побудувати толерантні відношення R_1 і R_2 на множині $M = \{1, 2, 3\}$, композиція $R_1 \circ R_2$ яких не є толерантним відношенням.

19. Довести, що коли R — рефлексивне і транзитивне відношення, тоді $R^{(k)} = R$ для всіх натуральних k.

20. Довести, що $R^+ = R^{(1)} \cup R^{(2)} \cup \ldots \cup R^{(k)} \cup \ldots$.

21. Довести, що відношення R транзитивне тоді й тільки тоді, коли $R^+ = R$.

22. Нехай C — матриця відношення R, заданого на скінченній множині M ($|M| = n$). Побудувати матрицю $C^{(k)}$ відношення $R^{(k)}$, $k = 0, 1, 2, \ldots$.

1.11. Відношення еквівалентності

Відношення R на множині M називається ***відношенням еквівалентності*** (або просто ***еквівалентністю***), якщо воно рефлексивне, симетричне і транзитивне.

Враховуючи важливість відношення еквівалентності, дамо розгорнуте означення цього поняття. Відношення R на множині M називають *відношенням еквівалентності*, або *еквівалентністю*, якщо воно має такі властивості:

1) $a\ R\ a$ для всіх $a\in M$ (рефлексивність),
2) якщо $a\ R\ b$, то $b\ R\ a$ для $a, b\in M$ (симетричність),
3) якщо $a\ R\ b$ і $b\ R\ c$, то $a\ R\ c$ для $a, b, c\in M$ (транзитивність).

Приклад 1.15. 1. Відношення рівності i_M на будь-якій множині M є відношенням еквівалентності. Рівність — це мінімальне відношення еквівалентності, бо з видаленням бодай одного елемента з i_M відношення перестає бути рефлексивним, а отже, і відношенням еквівалентності.

2. Відношення рівнопотужності множин є еквівалентністю (див. співвідношення (1.9)).

3. Важливу роль у математиці відіграє відношення "мають однакову остачу при діленні на k" або "конгруентні за модулем k", яке є відношенням еквівалентності на множині N натуральних чисел для будь-якого фіксованого $k\in N$. *Відношення конгруентності за модулем* k часто позначають $a \equiv b \pmod k$ (читається "a і b конгруентні за модулем k"). Цьому відношенню належать, наприклад, пари натуральних чисел $(17, 22)$, $(1221, 6)$, $(42, 57)$ для $k=5$, тобто $17 \equiv 22 \pmod 5$, $1221 \equiv 6 \pmod 5$, $42 \equiv 57 \pmod 5$.

4. Еквівалентністю є відношення подібності на множині всіх трикутників.

Сукупність множин $\{B_i \mid i\in I\}$ називається ***розбиттям*** множини A, якщо $\bigcup_{i\in I} B_i = A$ і $B_i \cap B_j = \varnothing$ для $i \neq j$. Множини B_i, $i\in I$, є підмножинами множини A. Їх називають ***класами, суміжними класами, блоками*** чи ***елементами розбиття***. Кожний елемент $a\in A$ належить одній і тільки одній множині B_i, $i\in I$.

Припустимо, що на множині M задано відношення еквівалентності R. Виконаємо таку побудову. Виберемо якийсь елемент $a\in M$ і утворимо підмножину $S_a^R=\{x \mid x\in M \text{ і } a\ R\ x\}$, що складається з усіх елементів множини M, еквівалентних елементу a. Відтак візьмемо другий елемент $b\in M$ такий, що $b\notin S_a^R$, й утворимо множину $S_b^R=\{x \mid x\in M \text{ і } b\ R\ x\}$ з елементів, еквівалентних b, і т. д. Таким способом одержимо сукупність множин (можливо, нескінченну) $\{S_a^R, S_b^R, \ldots\}$. Побудовану сукупність $\{S_i^R \mid i\in I\}$ називають ***фактор-множиною*** множини M за еквівалентністю R і позначають M/R.

Приклад 1.16. 1. Фактор-множина за відношенням рівності i_M для будь-якої множини M має вигляд $M/i_M = \{\{a\} \mid a\in M\}$.

2. Фактор-множина для відношення "конгруентні за модулем 3" на множині N натуральних чисел складається з трьох класів $\{3k \mid k\in N\}$, $\{3k - 1 \mid k\in N\}$ і $\{3k-2 \mid k\in N\}$.

Доведемо, що фактор-множина M/R є розбиттям множини M. Оскільки за побудовою кожний елемент множини M належить принаймні одній з множин S_i^R, $i\in I$, то $\bigcup_{i\in I} S_i^R = M$. Відтак припустимо, що для якихось $S_a^R \neq S_b^R$ існує елемент $c\in S_a^R \cap S_b^R$. Тоді з $c\in S_a^R$ випливає $a\ R\ c$, а з $c\in S_b^R$ — $b\ R\ c$. Із симетричності та транзитивності відношення R виводимо $a\ R\ b$ і $b\ R\ a$. Зі співвідношення $a\ R\ b$ і правила побудови множини S_a^R маємо $S_a^R \subseteq S_b^R$, а з $b\ R\ a$ і правила побудови множини S_b^R одержуємо протилежне включення $S_b^R \subseteq S_a^R$. Отже, $S_a^R = S_b^R$, і з одержаної суперечності випливає справедливість сформульованого твердження.

Будь-які два елементи з одного класу S_i^R еквівалентні між собою, а будь-які два елементи з різних класів фактормножини M/R нееквівалентні. Класи S_i^R називають ***класами еквівалентності*** за відношенням R. Клас еквівалентності, що містить елемент $a\in M$, часто позначають через $[a]_R$.

Потужність $|M/R|$ фактор-множини M/R називається ***індексом розбиття***, або ***індексом відношення еквівалентності*** R.

З іншого боку, припустимо, що для множини M задано якесь розбиття $\{S_i \mid i\in I\}$. Побудуємо відношення T на множині M за таким правилом: $a\ T\ b$ для $a, b\in M$ тоді й тільки тоді, коли елементи a і b належать одному класу даного розбиття. Неважко переконатися, що відношення T рефлексивне, симетричне і транзитивне, тобто є еквівалентністю на множині M. Отже, справедлива така теорема.

Теорема 1.10. Існує взаємно однозначна відповідність між усіма можливими еквівалентностями на множині M і всіма розбиттями цієї множини. Інакше кажучи, кожному відношенню еквівалентності на множині M відповідає єдине розбиття цієї множини на класи і, навпаки, кожне розбиття множини M однозначно задає певне відношення еквівалентності на M.

Нехай R — відношення еквівалентності на множині M. Відображення множини M на фактор-множину M/R, що кожному елементу $a \in M$ ставить у відповідність клас еквівалентності $[a]_R$, якому належить елемент a, називається ***канонічним***, або ***природним відображенням*** множини M на фактор-множину M/R.

1.10. Задачі і вправи

1. На множині $N \times N$ означимо відношення R і Q:

(а) $((a, b), (c, d)) \in R \quad \Leftrightarrow \quad a + d = b + c$;

(б) $((a, b), (c, d)) \in Q \quad \Leftrightarrow \quad ad = bc$.

Довести, що R і Q — відношення еквівалентності на множині $N \times N$.

2. Нехай $M = N \times N$. Означимо на множині M відношення R: $(a, b)\ R\ (c, d)$ тоді й тільки тоді, коли

(а) $ab = cd$; (б) $a + b = c + d$.

Довести, що R — еквівалентність на M. Виписати всі елементи класів еквівалентності $[(1, 1)]$, $[(2, 2)]$, $[(4, 3)]$, $[(1, 23)]$ та $[(6, 8)]$ за відношенням R.

3. На множині N натуральних чисел означимо відношення R: mRn тоді й тільки тоді, коли $m/n = 2^k$ для якогось цілого k.

(а) Довести, що R — відношення еквівалентності на N.

(б) Скільки різних класів еквівалентності є серед $[1]_R$, $[2]_R$, $[3]_R$ і $[4]_R$?

(в) Скільки різних класів еквівалентності є серед $[6]_R$, $[7]_R$, $[12]_R$, $[24]_R$, $[28]_R$, $[42]_R$ і $[48]_R$?

4. Нехай у множині M зафіксовано якусь підмножину $K \subseteq M$. Означимо відношення R на $\beta(M)$: $A\,R\,B$ тоді й тільки тоді, коли $A \cap K = B \cap K$, $A, B \in \beta(M)$.

(а) Довести, що R — відношення еквівалентності на $\beta(M)$.

(б) Для $M = \{1, 2, 3\}$ та $K = \{1, 2\}$ знайти класи еквівалентності за відношенням R.

(в) Для $M = \{1, 2, 3, 4, 5\}$ і $K = \{1, 2, 3\}$ визначити $[A]_R$, де $A = \{2, 3, 4\}$.

(г) Для $M = \{1, 2, 3, 4, 5\}$ і $K = \{1, 2, 3\}$ визначити кількість класів еквівалентності та кількість підмножин множини M у кожному класі еквівалентності за відношенням R.

(д) Визначити кількість класів еквівалентності та кількість підмножин множини M у кожному класі еквівалентності за відношенням R, якщо $|\,M\,| = n$ і $|\,K\,| = m$.

5. Довести, що перетин будь-якої сукупності відношень еквівалентності на множині M є еквівалентністю на M.

6. Навести приклад двох еквівалентностей R_1 і R_2 на множині $M = \{1, 2, 3\}$ таких, що $R_1 \cup R_2$ не є еквівалентністю на множині M.

7. Довести, що об'єднання $R_1 \cup R_2$ двох еквівалентностей R_1 і R_2 є еквівалентністю тоді й тільки тоді, коли $R_1 \cup R_2 = R_1 \circ R_2$.

8. Довести, що для довільного відношення еквівалентності R виконується рівність $R \circ R = R$.

9. Довести, що композиція $R_1 \circ R_2$ двох еквівалентностей R_1 і R_2 є еквівалентністю, якщо $R_1 \circ R_2 = R_1 \cup R_2$.

10. Нехай R_1 і R_2 — еквівалентності і $R_1 \circ R_2 = R_2 \circ R_1$. Довести, що $R_1 \circ R_2$ — найменше відношення еквівалентності, яке містить $R_1 \cup R_2$.

11. Побудувати найменше відношення еквівалентності Q на множині $M = \{1, 2, 3, 4\}$, що включає задане відношення R.

(а) $R = \{(2, 4), (3, 1)\}$; (б) $R = \{(2, 1), (1, 2), (2, 3), (3, 3)\}$.

12. Нехай R — відношення на M. Довести, що R є еквівалентністю тоді й тільки тоді, коли $(R \circ R^{-1}) \cup i_M = R$.

13. Довести, що для довільного відношення еквівалентності R наведені твердження рівносильні:

(а) $(x, y) \in R$; (б) $[x]_R \cap [y]_R \neq \varnothing$; (в) $[x]_R = [y]_R$.

14. Нехай $f\colon A \to B$ — довільне відображення і $R = \{(x, y) \mid x, y \in A, f(x) = f(y)\}$. Довести, що R — еквівалентність на множині A, а для відображення f існує розклад $f = \varepsilon \circ f_1$, де ε — природне відображення множини A на A/R, а f_1 — взаємно однозначна відповідність між A/R і $f(A)$ (*правило розкладу, або факторизації відображення*).

15. Нехай M — скінченна множина. Яке відношення еквівалентності на M має

(а) найбільший індекс;

(б) найменший індекс?

16. Нехай R — відношення еквівалентності на скінченній множині M ($|M| = n$) і $|R| = k$. Довести, що $k - n$ завжди парне число.

1.12. Відношення порядку

Відношення R на множині M називається ***відношенням часткового (нестрогого) порядку***, якщо воно рефлексивне, антисиметричне та транзитивне, тобто

1) $a\,R\,a$ для всіх $a \in M$ (рефлексивність),
2) якщо $a\,R\,b$ та $b\,R\,a$, то $a = b$ (антисиметричність),
3) якщо $a\,R\,b$ та $b\,R\,c$, то $a\,R\,c$ (транзитивність).

Множина M, на якій задано деякий частковий порядок, називається ***частково впорядкованою***.

Елементи a, $b \in M$ назвемо ***порівнюваними*** за відношенням R, якщо $a\,R\,b$ або $b\,R\,a$.

Частково впорядкована множина M, у якій будь-які два елементи порівнювані між собою, називається ***лінійно впорядкованою*** множиною або ***ланцюгом***. Відповідне відношення R, задане на лінійно впорядкованій множині, називається ***лінійним (досконалим) порядком***. Отже, відношення R на множині M називається відношенням *лінійного порядку*, якщо воно рефлексивне, антисиметричне, транзитивне і для будь-якої пари елементів a, $b \in M$ виконується $a\,R\,b$ або $b\,R\,a$.

Для позначення відношень порядку будемо використовувати звичайні математичні знаки $\leq$ і $\geq$. Інакше кажучи, для відношення порядку R замість $a\,R\,b$ будемо записувати $a \leq b$ або $b \geq a$ і читати відповідно "a менше або дорівнює b" або "b більше або дорівнює a". Очевидно, що $\leq$ — обернене відношення до $\geq$.

За кожним відношенням часткового порядку $\leq$ на довільній множині M можна побудувати інше відношення $<$ на M, вважаючи, що $a < b$ тоді й лише тоді, коли $a \leq b$ і $a \neq b$. Воно називається ***відношенням строгого порядку*** на множині M. Зрозуміло, що відношення строгого порядку антирефлексивне, транзитивне, а також задовольняє умову так званої сильної антисиметричності, чи асиметричності, тобто для жодної пари a, $b \in M$ не можуть одночасно виконуватися нерівності $a < b$ і $b < a$.

З іншого боку, за довільним відношенням строгого порядку $<$ на множині M можна однозначно побудувати відповідне відношення часткового (нестрогого) порядку $\leq$, поклавши $a \leq b$ тоді й тільки тоді, коли $a < b$ або $a = b$, a, $b \in M$. З огляду на такий простий зв'язок між відношеннями часткового (нестрогого) і строгого порядку можна обмежитися вивченням лише одного з них, наприклад $\leq$.

Приклад 1.17. 1. Традиційні відношення $\leq$ і $<$ ($\geq$ і $>$) — це відношення відповідно часткового і строгого порядку на множинах чисел N, Z і R. Більше того, множини N, Z і R, а також будь-які їх підмножини лінійно впорядковані за відношеннями $\leq$ або $\geq$.

2. Частковим порядком є відношення рівності i_M на будь-якій множині M. Цей порядок іноді називають ***тривіальним***.

3. Відношення нестрогого включення $\subseteq$ є відношенням часткового порядку, а відношення $\subset$ — відношенням строгого порядку на множині $\beta(A)$ всіх підмножин (булеані) заданої множини A.

4. Задамо відношення $\leq$ і $<$ на множині R^n кортежів дійсних чисел довжини n так: $(a_1, a_2, \ldots, a_n) \leq (b_1, b_2, \ldots, b_n)$ тоді й тільки тоді, коли $a_1 \leq b_1$, $a_2 \leq b_2$, …, $a_n \leq b_n$; аналогічно $(a_1, a_2, \ldots, a_n) < (b_1, b_2, \ldots, b_n)$ тоді й лише тоді, коли $(a_1, a_2, \ldots, a_n) \leq (b_1, b_2, \ldots, b_n)$ і принаймні для однієї координати $i = 1, 2, \ldots, n$ виконується $a_i < b_i$. Тоді $(2, 3.7, 4) < (7, 24, 10)$, а кортежі $(1, 4, -1.7)$ і $(2, 2, 4)$ непорівнювані. Аналогічно можна ввести частковий порядок на множинах N^n, Z^n і Q^n.

5. Зафіксуємо строгий порядок розміщення символів у довільному скінченному алфавіті $A = \{a_1, a_2, \ldots, a_n\}$, наприклад, покладемо, що $a_1 < a_2 < \ldots < a_n$. Тоді природним чином можна означити так званий ***лексикографічний порядок*** на множині A^m всіх слів довжини m в алфавіті A. А саме, вважатимемо $a_{i_1}a_{i_2}\ldots a_{i_m} < a_{j_1}a_{j_2}\ldots a_{j_m}$ тоді й тільки тоді, коли $a_{i_s} = a_{j_s}$ для $s = 1, 2, \ldots, k-1$ і $a_{i_k} < a_{j_k}$ для певного k, $k = 1, 2, \ldots, m$.

Лексикографічний порядок можна поширити на множину A^* всіх слів в алфавіті A, якщо доповнити його додатковим ("порожнім") символом p і вважати, що $p < a_i$ для всіх $i = 1, 2, \ldots, n$. При порівнюванні двох слів різної довжини спочатку слово меншої довжини доповнюють справа такою кількістю "порожніх" символів p, щоб воно зрівнялося за довжиною з другим словом, після чого обидва слова порівнюють як слова однакової довжини.

Нехай $A = \{a, b, c\}$ і $a < b < c$, тоді $aac < aba$, $ab < abab$, $b < cba$ тощо.

Лексикографічний порядок лежить в основі впорядкування всіх словників, енциклопедій, індексів (предметних або іменних покажчиків), довідників, списків, таблиць тощо.

6. У множині N натуральних чисел відношення "ділить" — це частковий порядок. Маємо $4 \leq 28$, $11 \leq 132$, $5 \leq 5$, $1 \leq n$ для будь-якого $n \in N$, а пари чисел 7 і 22, 13 і 35 непорівнювані.

Нехай M — частково впорядкована множина, A — деяка її непорожня підмножина. ***Верхньою гранню*** підмножини $A \subseteq M$ у множині M називається такий елемент $b \in M$, що $a \leq b$ для всіх $a \in A$. Елемент b називається ***найбільшим елементом*** множини M, якщо b — верхня грань множини M.

Аналогічно, елемент c частково впорядкованої множини M називається ***нижньою гранню*** підмножини $A \subseteq M$, якщо $c \leq a$ для будь-якого $a \in A$. Елемент c — ***найменший*** у множині M, якщо c — нижня грань самої множини M.

Отже, найбільший і найменший елементи, а також верхня та нижня грані (якщо вони існують) порівнювані відповідно з усіма елементами множини M або підмножини A.

Елемент $x \in M$ називається ***максимальним*** у множині M, якщо не існує такого елемента $a \in M$, що $x < a$. Відповідно елемент $y \in M$ називається ***мінімальним*** у множині M, якщо не існує такого елемента $a \in M$, що $a < y$.

Очевидно, що коли в частково впорядованій множині M існує найбільший елемент, то це єдиний максимальний елемент множини M. Аналогічно, найменший елемент множини M — єдиний її мінімальний елемент. Зауважимо також, що частково впорядкована множина M може не мати найбільшого (найменшого) елемента й водночас мати один або кілька максимальних (міиімальних) елементів. У лінійно впорядкованій множині поняття найбільшого та максимального (найменшого і мінімального) елементів збігаються.

Приклад 1.18. 1. У множині Z цілих чисел з традиційним відношенням порядку множина N натуральних чисел має найменший елемент (число 1) і не має найбільшого елемента. Будь-яке невід'ємне число, а також 1 є нижніми гранями для N.

2. У довільній множині M з тривіальним порядком i_M (відношенням рівності) кожен елемент $a \in M$ є одночасно максимальним і мінімальним. Якщо $|M| > 1$, то найбільшого і найменшого елементів у множині M немає.

3. Булеан $\beta(A)$ множини A з відношенням часткового порядку $\subseteq$ містить найменший елемент — порожню множину і найбільший елемент — саму множину A. У множині D всіх непорожніх підмножин множини A (тобто у множині $\beta(A) \backslash \{\varnothing\}$) немає найменшого елемента, але всі одноелементні множини $\{a\}$, $a \in A$, є мінімальними елементами множини D.

4. У множині M усіх натуральних дільників числа $n \in N$, частково впорядкованій за відношенням "ділить", число 1 — найменший елемент, а n — найбільший.

Лінійно впорядкована множина (ланцюг) M називається ***цілком упорядкованою***, якщо кожна її непорожня підмножина $A \subseteq M$ має найменший елемент.

Зокрема, множина N натуральних чисел з традиційним відношенням порядку цілком упорядкована, а множина Z цілих чисел не є цілком упорядкованою, оскільки будь-яка її нескінченна підмножина від'ємних чисел не має найменшого елемента.

Якщо M — частково впорядкована множина, то множина L усіх її ланцюгів (тобто лінійно впорядкованих підмножин) також частково впорядкована за відношенням теоретико-множинного включення. Максимальні елементи множини L називають ***максимальними ланцюгами*** множини M.

Наведемо низку важливих тверджень про частково впорядковані множини, які часто застосовують у вищій алгебрі, математичному аналізі, топології та інших розділах сучасної математики.

Теорема Куратовського–Цорна, або ***лема Цорна***. Якщо в частково впорядкованій множині M будь-який ланцюг має верхню (нижню) грань, то множина M має максимальний (мінімальний) елемент.

Теорема Хаусдорфа. Будь-який ланцюг частково впорядкованої множини M міститься в деякому максимальному ланцюзі множини M.

Теорема Цермело. Будь-яку множину M можна цілком упорядкувати.

Можна довести, що всі три наведені теореми рівносильні між собою (тобто зі справедливості будь-якої з них випливає справедливість двох інших), і що всі вони рівносильні одній із фундаментальних аксіом сучасної аксіоматичної теорії множин — так званій *аксіомі вибору* [1, 19, 32].

Аксіома вибору. Якщо дано множину M, то існує функція $w\colon (\beta(M) \backslash \{\varnothing\}) \to M$, яка кожній непорожній підмножині $A \subseteq M$ ставить у відповідність певний елемент $a = w(A)$ цієї підмножини, $a \in A$.

Інакше кажучи, функція w обирає по одному елементу з непорожніх підмножин множини M.

Зокрема, ми неявно користувалися аксіомою вибору у доведенні теореми 1.2.

Будь-яке твердження T стосовно елементів деякої цілком упорядкованої множини M можна доводити, використовуючи так звану ***трансфінітну індукцію*** — узагальнення відомого методу математичної індукції.

База індукції. Доводимо виконання твердження T для найменшого елемента множини M.

Індукційний крок. Робимо припущення, що твердження T виконується для будь-якого елемента a такого, що $a < b$, і доводимо, що твердження T справджується для елемента b.

Якщо виконуються умови бази й індукційного кроку, то твердження T справджується для довільного елемента $a \in M$.

Припустімо, що це не так. Нехай у множині M існують елементи, для яких не виконується твердження T, і $K \subseteq M$ — сукупність усіх таких елементів. Множина M цілком упорядкована, отже K має найменший елемент $b \in K$. З умови бази індукції випливає, що b не є найменшим елементом у множині M. Це означає, що у множині M існують елементи $a < b$, і для всіх таких елементів a виконується твердження T. Згідно з індукційним кроком твердження T має виконуватись і для елемента b, що призводить до суперечності.

1.11. Задачі і вправи

1. Означимо відношення R на множині цілих чисел Z так: $m\,R\,n$ тоді й тільки тоді, коли $m - n$ — невід'ємне парне число. Довести, що R — частковий порядок на Z. Чи є R лінійним порядком?

2. Нехай M — довільна множина. Означимо відношення R на множині $\beta(M) \times \beta(M)$: $(A, B)\,R\,(C, D)$ тоді й тільки тоді, коли $A \Delta B \subseteq C \Delta D$, $A, B, C, D \in \beta(M)$. Чи є R відношенням часткового порядку?

3. Нехай $\le_A$ — частковий порядок на множині A, $\le_B$ — частковий порядок на множині B. Назвемо *прямим добутком* частково впорядкованих множин A і B множину $A \times B$ із заданим на ній відношенням $\le$: $(a_1, b_1) \le (a_2, b_2) \Leftrightarrow$ $\Leftrightarrow a_1 \le_A a_2$ і $b_1 \le_B b_2$. Довести, що $\le$ — частковий порядок на $A \times B$.

4. Довести чи спростувати таке твердження: якщо $\le_A$ — лінійний порядок на множині A, а $\le_B$ — лінійний порядок на множині B, то відношення $\le$, означене в попередній задачі, є лінійним порядком на множині $A \times B$.

5. Нехай M — довільна множина. Означимо відношення R на множині $\beta(M) \times \beta(M)$: $(A, B)\,R\,(C, D)$ тоді й тільки тоді, коли $A \subseteq C$ і $B \subseteq D$, $A, B, C, D \in \beta(M)$. Визначити, чи є R:

(а) відношенням часткового порядку;

(б) відношенням лінійного порядку.

6. Нехай R — транзитивне відношення на множині M. Довести, що R є частковим порядком на M тоді й тільки тоді, коли $R \cap R^{-1} = i_M$.

7. Довести, що об'єднання $R_1 \cup R_2$ відношень часткового порядку R_1 і R_2 на множині M є частковим порядком на множині M тоді й тільки тоді, коли $R_1 \circ R_2 \cup R_2 \circ R_1 \subseteq R_1 \cup R_2$ і $R_1 \cap R_2^{-1} = i_M$.

8. Довести, що для довільної частково впорядкованої множини M з k елементами існує таке відображення f: $M \to N_k$, що для елементів $a_i, a_j \in M$ зі співвідношення $a_i < a_j$ випливає нерівність $f(a_i) < f(a_j)$.

9. Нехай $\le$ і $<$ — традиційні відношення порядку на множині натуральних чисел N. Довести, що

(а) $< \circ < \ne <$; (б) $\le \circ < = <$; (в) $\le \circ \ge = N^2$.

10. Довести, що будь-яка непорожня скінченна частково впорядкована множина A містить мінімальний і максимальний елементи.

11. Довести, що скінченна частково впорядкована множина має найменший (найбільший) елемент тоді й тільки тоді, коли вона містить рівно один мінімальний (максимальний) елемент. Чи це так для нескінченних частково впорядкованих множин?

12. Чи є цілком упорядкованою

(а) множина Q раціональних чисел з традиційним відношенням порядку;

(б) множина R дійсних чисел з традиційним відношенням порядку;

(в) множина чисел виду $1 - 1/n$, де $n \in N$, з традиційним відношенням порядку?

13. Знайти всі множини M, для яких існує повний порядок R на M такий, що R^{-1} є повним порядком на M.

14. Нехай A — цілком упорядкована множина. Довести, що не існує такого монотонного взаємно однозначного відображення f: $A \to A$, що для деякого елемента $a \in A$ виконується $f(a) < a$. (Відображення f частково впорядкованої множини A у частково впорядковану множину B називається *монотонним*, якщо з $a < b$ випливає $f(a) < f(b)$ для всіх порівнюваних елементів $a, b \in A$.)

1.13. Решітки

Серед частково впорядкованих множин винятково важливу роль відіграють так звані решітки, або структури.

Точною верхньою гранню підмножини A частково впорядкованої множини M (позначають $\sup A$; читають "супремум A") називають найменший елемент у множині всіх верхніх граней підмножини A. Відповідно, ***точною нижньою гранню*** підмножини A частково впорядкованої множини M (позначають $\inf A$; читають "інфімум A") називають найбільший елемент у множині всіх нижніх граней підмножини A.

Частково впорядкована множина M називається ***решіткою*** (***структурою*** або ***ґраткою***), якщо для будь-якої пари елементів $a, b \in M$ (тобто для будь-якої двоелементної підмножини множини M) існують $\sup\{a, b\}$ та $\inf\{a, b\}$.

Приклад 1.19. 1. Будь-яка лінійно впорядкована множина M (наприклад, числові множини N, Z, Q і R з традиційними відношеннями порядку) є решіткою. Якщо $a, b \in M$, то $\sup\{a, b\} = \max(a, b)$ й $\inf\{a, b\} = \min(a, b)$.

2. Розглянемо множину N натуральних чисел з відношенням часткового порядку “ділить”. Для довільних a, $b \in N$ маємо $\sup\{a, b\} = $ НСК(a, b) й $\inf\{a, b\} = $ НСД(a, b) (НСК — найменше спільне кратне, НСД — найбільший спільний дільник). Наприклад, $\sup\{12, 32\} = 96$, $\inf\{12, 32\} = 4$, $\inf\{16, 27\} = 1$ тощо.

3. Частково впорядкована за відношенням включення множина $\beta(M)$ усіх підмножин множини M є решіткою: $\sup\{A, B\} = A \cup B$ й $\inf\{A, B\} = A \cap B$, $A, B \subseteq M$.

4. Розглянемо множину R^n кортежів дійсних чисел довжини n з відношенням часткового порядку, означеним у прикладі 1.17(4), тобто $(a_1, a_2, \ldots, a_n) \le (b_1, b_2, \ldots, b_n)$ тоді й тільки тоді, коли $a_i \le b_i$, $i = 1, 2, \ldots, n$. Частково впорядкована у такий спосіб множина R^n є решіткою:

$$\sup\{(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n)\} = (c_1, c_2, \ldots, c_n),$$

де $c_i = \max(a_i, b_i)$, $i = 1, 2, \ldots, n$,

$$\text{а } \inf\{(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n)\} = (d_1, d_2, \ldots, d_n),$$

де $d_i = \min(a_i, b_i)$, $i = 1, 2, \ldots, n$.

Аналогічно можна перетворити на решітки множини кортежів N^n, Z^n, Q^n і B^n, де $B = \{0, 1\}$ — множина двійкових цифр.

5. Множину $P = \{R_1, R_2, \ldots, R_m\}$ усіх можливих розбиттів певної скінченної множини M можна перетворити в решітку в такий спосіб. Вважатимемо, що розбиття $R_i = \{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\}$ і $R_j = \{A_{j_1}, A_{j_2}, \ldots, A_{j_l}\}$ перебувають у відношенні $R_i \le R_j$, якщо кожен клас A_{i_t}, $t = 1, 2, \ldots, k$, розбиття R_i міститься в деякому класі A_{j_s} розбиття R_j. Наприклад, для $M = \{1, 2, 3, 4, 5\}$ розбиття $R' = \{\{1, 2\}, \{3\}, \{4, 5\}\}$ менше від розбиття $R'' = \{\{1, 2, 3\}, \{4, 5\}\}$ і менше від розбиття $R''' = \{\{1, 2\}, \{3, 4, 5\}\}$, а розбиття R'' і R''' непорівнювані.

Найменшим елементом частково впорядкованої множини P є розбиття $\{\{a\} \mid a \in M\}$, а найбільшим — $\{M\}$. Тоді $\sup\{R_i, R_j\} = R_k$, де R_k — розбиття, у якому елементи a, $b \in M$ належать одному класу тоді й тільки тоді, коли існує такий елемент $c \in M$, що кожна з пар елементів a і c та c і b належить одному класу або в R_i, або в R_j; $\inf\{R_i, R_j\} = R_l$, де R_l — розбиття, в якому елементи a, $b \in M$ належать одному класу тоді й тільки тоді, коли вони належать одному класу і в R_i, і в R_j.

Наприклад,

$\sup\{R'', R'''\} = \{\{1, 2, 3, 4, 5\}\}$, $\inf\{R'', R'''\} = \{\{1, 2\}, \{3\}, \{4, 5\}\}$,

$\sup\{R', R''\} = \{\{1, 2, 3\}, \{4, 5\}\}$, $\inf\{R', R''\} = \{\{1, 2\}, \{3\}, \{4, 5\}\}$.

Оскільки за теоремою 1.10 існує взаємно однозначна відповідність між усіма розбиттями множини M і всіма відношеннями еквівалентнос-

ті на M, то множину всіх відношень еквівалентності на M також можна перетворити в решітку.

Скінченну частково впорядковану множину M найчастіше зображають за допомогою *спрощеної діаграми* (графа) або так званої *діаграми Гессе* [2]. Для зменшення кількості стрілок з вершини a проводять стрілку у вершину b тільки тоді, коли $a \le b$ і не існує такого елемента c, що $a \le c$ і $c \le b$, а стрілки (петлі), що відповідають діагональним парам (a, a), взагалі не проводять.

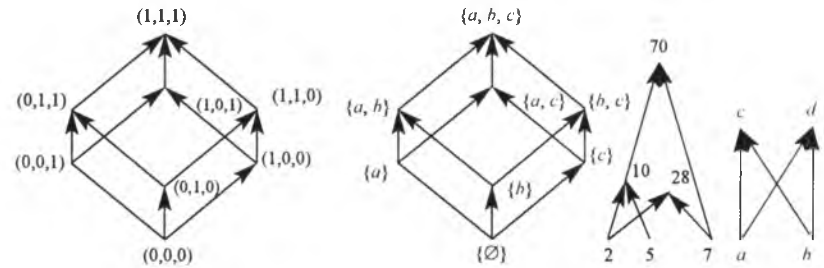
Приклад 1.20. 1. Розглянемо спрощені діаграми відповідно для чотирьох частково впорядкованих множин (рис. 1.5):

а) множини двійкових кортежів B^3;

б) булеана $\beta(M)$ множини $M = \{a, b, c\}$ з відношенням включення $\subseteq$;

в) множини натуральних чисел $C = \{2, 5, 7, 10, 28, 70\}$ з відношенням “ділить”;

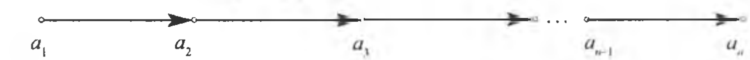
г) множини $D = \{a, b, c, d\}$ з відношенням часткового порядку $R = \{(a, a), (b, b), (c, c), (d, d), (a, c), (b, c), (a, d), (b, d)\}$.



а б в г

Рис. 1.5

2. Діаграма будь-якої скінченної лінійно впорядкованої множини $M = \{a_1, a_2, \ldots, a_n\}$, $a_i < a_{i+1}$, $i = 1, 2, \ldots, n-1$, має вигляд



Неважко переконатися, що $a \le b$, $a, b \in M$, тоді й тільки тоді, коли в такій діаграмі частково впорядкованої множини M існує складений зі стрілок шлях, що веде з вершини a у вершину b. Верхня грань для $\{a, b\}$ — це елемент, у який ведуть шляхи з a і з b. Нижня грань $\{a, b\}$ — це елемент, з якого є шляхи і в a, і в b.

Частково впорядкована множина не є решіткою, якщо:

1) якась пара елементів не має верхньої або нижньої грані;

2) для деякої пари елементів не існує найменшої верхньої (або найбільшої нижньої) грані.

Наприклад, перші дві множини B і $\beta(M)$ із прикладу 1.20 є решітками, тому що для їх діаграм не виконується жодна з наведених умов. Множина C не є решіткою, бо, наприклад, для пар $\{2, 5\}$, $\{5, 7\}$, $\{7, 10\}$ не існує нижніх граней, а пари $\{10, 28\}$ і $\{28, 70\}$ не мають верхніх граней. Пара елементів $\{a, b\}$ $(\{c, d\})$ множини D має дві верхні (дві нижні) грані c і d (відповідно a і b), однак не має найменшої верхньої (найбільшої нижньої) грані, оскільки елементи c і d (a і b) непорівнювані між собою.

Частково впорядкована множина M називається ***повною решіткою***, якщо для будь-якої непорожньої підмножини $A \subseteq M$ у множині M існує найменша верхня грань $\sup A$ і найбільша нижня грань $\inf A$.

Очевидно, що довільна повна решітка є решіткою, але не будь-яка решітка є повною решіткою. Якщо M — повна решітка, то найменшу верхню грань усієї множини M ($\sup M$) називають ***одиницею*** цієї решітки і позначають **1**, а найбільшу нижню грань множини M ($\inf M$) — ***нулем*** решітки і позначають **0**. Вибір цих назв для $\sup M$ та $\inf M$ пояснюється такими властивостями елементів **1** і **0**. Для довільного елемента $a \in M$ виконуються такі умови:

$$\sup\{\mathbf{1}, a\} = \mathbf{1}, \quad \sup\{\mathbf{0}, a\} = a, \quad a \le \mathbf{1},$$
$$\inf\{\mathbf{1}, a\} = a, \quad \inf\{\mathbf{0}, a\} = \mathbf{0}, \quad \mathbf{0} \le a.$$

Елементи **0** і **1** є відповідно найменшим і найбільшим елементами повної решітки M.

Приклад 1.21. 1. Решітки B^3, $\beta(M)$ і P з прикладу 1.19 — повні. Одиницями цих решіток є відповідно (1, 1, 1), M і $\{M\}$, а нулями — (0, 0, 0), $\varnothing$ і $\{\{a\} \mid a \in M\}$.

2. Множина N натуральних чисел з традиційним відношенням порядку є решіткою, але не є повною решіткою, оскільки будь-яка її нескінченна підмножина не має найменшої верхньої грані.

3. Множина всіх дільників натурального числа n, частково впорядкована за відношенням "ділить", є повною решіткою. Одиницею в такій решітці є число n, а нулем — число 1.

1.12. Задачі і вправи

1. Сформулювати умови для діаграми частково впорядкованої множини, при виконанні або порушенні яких ця множина є чи не є решіткою.

2. Довести, що в решітці будь-який максимальний елемент є найбільшим, а будь-який мінімальний елемент — найменшим.

3. Довести, що в будь-якій скінченній решітці є найбільший і найменший елементи.

4. Довести, що в будь-якій решітці L для довільних елементів $a, b, c, d \in L$ виконується:

(а) $\sup(a, a) = a$, $\inf(a, a) = a$;

(б) $a \le \sup(a, b)$, $\inf(a, b) \le a$;

(в) якщо $a \le c$ і $b \le c$, то $\sup(a, b) \le c$;

(г) якщо $c \le a$ і $c \le b$, то $c \le \inf(a, b)$;

(д) $a \le b$ тоді й тільки тоді, коли $\sup(a, b) = b$;

(е) $a \le b$ тоді й тільки тоді, коли $\inf(a, b) = a$;

(є) якщо $a \le c$ і $b \le d$, то $\sup(a, b) \le \sup(c, d)$;

(ж) якщо $a \le c$ і $b \le d$, то $\inf(a, b) \le \inf(c, d)$;

(з) якщо $a \le c$, то $\sup(a, \inf(b, c)) \le \inf(\sup(a, b), c)$;

(и) $\sup(a, b) = \inf(a, b)$ тоді й тільки тоді, коли $a = b$.

5. Довести, що в будь-якій решітці L для довільних елементів $a, b, c \in L$ виконується:

(а) $\sup(a, b) = \sup(b, a)$, $\inf(a, b) = \inf(b, a)$ (*комутативність*);

(б) $\sup(a, \sup(b, c)) = \sup(\sup(a, b), c)$, $\inf(a, \inf(b, c)) = \inf(\inf(a, b), c)$ (*асоціативність*);

(в) $\sup(\inf(a, b), a) = a$, $\inf(\sup(a, b), a) = a$ (*закони поглинання*).

6. Довести, що частково впорядкована множина L є решіткою тоді й тільки тоді, коли для довільної скінченної непорожньої підмножини $A \subseteq L$ існують $\sup A$ та $\inf A$.

7. Чи утворює повну решітку впорядкована за відношенням включення множина:

(а) усіх рефлексивних;

(б) усіх антирефлексивних;

(в) усіх симетричних;

(г) усіх антисиметричних;

(д) усіх транзитивних;

(е) усіх толерантних

відношень на даній множині M?

8. Чи є повною решіткою множина чисел:

(а) $1 - 1/n$, $n \in N$;

(б) $1/n$, $n \in N$;

(в) усіх раціональних чисел зі сегмента [0, 1]?

9. Чи є повною решіткою множина A^* всіх слів у алфавіті A з лексико-графічним порядком?

10. Довести, що множина всіх дільників натурального числа n, частково впорядкована за відношенням "ділить", є повною решіткою. Визначити елементи **0** і **1** цієї решітки.

11. Довести, що будь-яка скінченна решітка є повною решіткою.

12. Довести, що в будь-якій решітці L для довільних елементів a, b, $c \in L$ виконується:

(а) $\inf(a, \sup(b, c)) \geq \sup(\inf(a, b), \inf(a, c))$;

(б) $\sup(a, \inf(b, c)) \leq \inf(\sup(a, b), \sup(a, c))$

(*закони напівдистрибутивності*).

1.14. Парадокси теорії множин

Слово "*парадокс*" грецького походження; українською мовою його перекладають як несподіваний, дивний. Це слово вживають щодо висловлювання (положення, ідеї), яке суттєво різниться від загальноприйнятого традиційного уявлення про це. Уживання терміна "парадокс" стосовно суперечностей, виявлених різними математиками в теорії множин Г. Кантора, є наївною спробою зменшити їх значення й надати їм характеру логічних курйозів, штучних, неприродних конструкцій. Точніше суть явища передає назва "*антиномії* теорії множин", оскільки термін антиномія є синонімом терміна "суперечність". Але за традицією будемо називати сформульовані нижче положення парадоксами.

Парадокс Б. Рассела. Для будь-якої множини M коректним є таке питання: чи множина M належить собі як окремий елемент, тобто чи є множина M елементом самої себе, чи ні? Наприклад, множина всіх множин є множиною, і тому належить сама собі, а множина всіх будинків у місті не є будинком, множина студентів в аудиторії не є студентом.

Отже, коректно поставити сформульоване питання й щодо множини всіх множин, які не є власними елементами. Нехай M — множина всіх тих множин, що не є елементами самих себе. Розглянемо питання: чи сама множина M є елементом самої себе чи ні? Якщо припустити, що $M \in M$, то з означення множини M випливає $M \notin M$. Якщо ж припустимо, що $M \notin M$, то з того самого означення дістанемо $M \in M$.

Близьким до парадокса Рассела є так званий "*парадокс цирульника*": цирульник — це мешканець міста, який голить тих і тільки тих мешканців міста, які не голять самі себе. Виникає проблема з визначенням множини C всіх тих мешканців міста, яких голить цирульник. Міркуючи

аналогічно тому, як це було зроблено в парадоксі Рассела, дійдемо висновку, що цирульник голить себе в тому й тільки в тому разі, коли він не голить сам себе, тобто цирульник належить множині C тоді й лише тоді, коли він не належить C.

А ось парадокс, що був відомий самому авторові теорії множин Г. Кантору. Розглянемо об'єднання всіх мислимих множин і позначимо його U. Тоді за теоремою 1.8 потужність множини $\beta(U)$ всіх підмножин множини U більша, ніж потужність самої множини U. Однак це парадоксально, бо за означенням множина U є множиною, що містить усі множини (зокрема, й множину $\beta(U)$).

Багато хто з математиків на початку XX століття не надавали цим парадоксам особливого значення, оскільки в той час теорія множин була відносно новою галуззю математики й не зачіпала інтересів більшості математиків. Однак їхні відповідальніші та проникливіші колеги зрозуміли, що виявлені парадокси стосуються не тільки теорії множин і побудованих на ній розділів класичної математики, але мають безпосереднє відношення до логіки взагалі як головного інструменту математики.

Зокрема, парадокс Рассела можна переформулювати в термінах логіки і таким чином додати до відомих з давніх часів логічних парадоксів: парадокса брехуна (людина, яка завжди каже неправду, одного разу мовить: "Те, що я сказав, — брехня"), парадокса всемогутньої істоти (чи *може* всемогутня істота створити такий камінь, який вона *не зможе* підняти?) тощо.

Гостро постало питання про обґрунтування засад математики. На початку XX століття виникли три основні напрями досліджень з обґрунтування сучасної математики. Коротко подамо суть кожного з них.

1. ***Логіцизм.*** Основною тезою логіцизму є положення про те, що першооснова математики — це логіка, а математика — лише її частина, тобто всі математичні істини складають власну підмножину множини всіх логічних істин.

Основні ідеї та методи логіцизму було вперше викладено у великій праці А. Уайтхеда та Б. Рассела "Принципи математики", що вийшла на початку другого десятиріччя XX століття.

Незважаючи на те, що в рамках логіцизму проблему обґрунтування математики не було остаточно розв'язано, все ж було чимало зроблено для з'ясування деяких важливих аспектів логічної структури математики.

2. ***Інтуїціонізм***. Основними засадами інтуїціонізму є такі принципи:

1) в основу математики кладуть поняття натурального числа, причому систему натуральних чисел вважають інтуїтивно відомою;

2) усі інші математичні об'єкти будують на основі натуральних чисел суто конструктивно за допомогою скінченної кількості застосувань скінченної кількості конкретних операцій. Доведення існування математичного об'єкта зводиться до побудови конкретного алгоритму, тобто визнаються лише конструктивні доведення існування математичних об'єктів. Зокрема, не визнається доведення існування математичних об'єктів методом від супротивного;

3) закон виключеного третього незастосовний до нескінченних множин (зазначений закон — це логічна аксіома, згідно з якою з двох тверджень "A" та "не A" тільки одне істинне);

4) визнається абстракція потенційної нескінченності та відкидається абстракція актуальної нескінченності.

Обґрунтування математики в межах інтуїціонізму натрапляє на дві основні перешкоди: значну частину важливих розділів математики не вдається побудувати засобами інтуїціонізму або ж така побудова має досить громіздкий і штучний вигляд, який не задовольняє переважну більшість як математиків-теоретиків, так і практиків.

Подальшим кроком у розвитку інтуїціонізму є конструктивний напрям (або *конструктивізм*), що розвивається на основі уточненого поняття алгоритму.

3. ***Формалізм***. Засновником формалізму вважають Д. Гільберта. Цей напрям є подальшим поглибленням аксіоматичного методу в математиці. Основою будь-якої аксіоматичної теорії є список неозначуваних (первинних) понять і список аксіом, тобто положень, які вважають вихідними й істинність яких декларують із самого початку. Додатково означають логічні правила, за допомогою яких з одних тверджень (зокрема, з аксіом) дістають інші.

Гільберт і його послідовники вважали, що кожен розділ математики можна повністю формалізувати, тобто за допомогою формальних виразів (формул) подати всі аксіоми, а всі математичні (логічні) доведення звести до суто формальних перетворень над формулами.

Саме на основі ідей формалізму Е. Цермело в 1908 р. побудував першу формальну аксіоматичну теорію множин (так звану *систему Цермело — Френкеля*, або *ZF* [19; 20; 32]). Пізніше було запропоновано багато видозмін і вдосконалень *ZF* і декілька інших аксіоматичних теорій множин.

Проаналізувавши всі парадокси теорії множин, можна дійти висновку, що всі вони зумовлені необмеженим застосуванням так званого *принципу абстракції* (або *принципу згортання*), згідно з яким для будь-якої властивості $P(x)$ існує відповідна множина елементів x, які мають властивість P. Якщо відкинути це припущення, то всі відомі парадокси теорії множин стають неможливими. Так, із парадоксів Рассела і Кантора випливало б, що не існують множина множин, які не є елементами самих себе, і множина всіх множин.

В усіх аксіоматичних теоріях множин неможливість антиномій ґрунтується на обмеженнях принципу згортання, тобто на обмеженні допустимих множин.

РОЗДІЛ 2

АЛГЕБРИЧНІ СИСТЕМИ

Математичні методи, поняття, основні об'єкти і результати алгебри (або, як часто її називають, *загальної алгебри*) широко використовують у різних розділах сучасної інформатики, зокрема в автоматизації проектування ЕОМ, теорії та практиці програмування, теорії автоматів, теорії формальних мов і граматик та ін.

2.1. Поняття алгебричної системи

Будь-яку функцію φ: $M^n \to M$ називають ***n*-арною операцією** на множині M; n називається ***арністю*** операції φ. Якщо $n = 2$, операція називається ***бінарною***, а для $n = 1$ — ***унарною***. Часто розглядають також і ***нульарні*** операції ($n = 0$). Кожна нульарна операція ψ фіксує чи виділяє у множині M деякий елемент $a \in M$, який має певні специфічні властивості, що відрізняють його від інших "звичайних" елементів цієї множини. Отже, для нульарної операції ψ маємо $\mathrm{Pr}_1\psi = \varnothing$ та $\mathrm{Pr}_2\psi = \{a\}$.

Елементи кортежу $\alpha = (a_1, a_2, \ldots, a_n) \in \mathrm{Pr}_1\varphi \subseteq M^n$ традиційно називають ***операндами***, або ***аргументами***, n-арної операції φ, а $\varphi(\alpha)$ — результатом дії, або просто ***результатом***, операції φ.

Множина M разом із заданими на ній сукупностями операцій $\Omega = \{\varphi_1, \varphi_2, \ldots, \varphi_m, \ldots\}$ і відношень $\pi = \{R_1, R_2, \ldots, R_k, \ldots\}$, тобто система $A = \langle M, \Omega, \pi\rangle$ називається ***алгебричною системою***.

Множина M називається ***носієм***, або ***основною множиною***, а пара множин $\langle\Omega, \pi\rangle$ — ***сигнатурою*** алгебричної системи A. Якщо обидві множини Ω та π — скінченні, то A називають *алгебричною системою скінченної сигнатури*.

Кортеж $(n_1, n_2, \ldots, n_m, \ldots, l_1, l_2, \ldots, l_k, \ldots)$ арностей операцій і відношень сигнатури називають ***типом*** алгебричної системи A.

Приклад 2.1. 1. На множині M можна побудувати багато різних алгебричних систем із тим самим носієм M і різними сигнатурами.

Наприклад, якщо N — множина натуральних чисел, то можна розглянути такі алгебричні системи: $A_1 = \langle N, \{+, \times\}, \varnothing\rangle$, $A_2 = \langle N, \{+, \times, 1\}, \{R_2, \le\}\rangle$, $A_3 = \langle N, \varnothing, \{D\}\rangle$. Тут $+$ і $\times$ — традиційні бінарні операції додавання і множення, 1 — нульарна операція, фіксований (виділений) елемент "одиниця", R_2 — унарне відношення "парність", $\le$ — бінарне відношення "менше або дорівнює", D — бінарне відношення "ділить". Типами алгебричних систем A_1, A_2 та A_3 є відповідно (2, 2), (2, 2, 0, 1, 2) і (2).

2. Одну з найпоширеніших у математиці алгебричних систем можна побудувати на основі розглянутої вище *решітки*. Нехай частково впорядкована множина M є решіткою. Означимо на множині M дві бінарні операції об'єднання $\cup$ і перетину $\cap$ для елементів $a, b \in M$ таким чином: $a \cup b = \sup\{a, b\}$ і $a \cap b = \inf\{a, b\}$. Тоді алгебрична система $P = \langle M, \{\cup, \cap\}, \{\le\}\rangle$ називається ***решіткою***. Її тип — (2, 2, 2).

Якщо в алгебричній системі $A = \langle M, \Omega, \pi\rangle$ множина операцій порожня ($\Omega = \varnothing$), то A називають ***моделлю*** і позначають $A = \langle M, \pi\rangle$. Якщо в A множина відношень порожня ($\pi = \varnothing$), то алгебричну систему A називають ***алгеброю*** і позначають $A = \langle M, \Omega\rangle$. Алгебрична система A_1 з останнього прикладу є алгеброю, а A_3 — моделлю.

Оскільки будь-яку n-арну операцію на множині M можна розглядати як $(n + 1)$-арне відношення на M, то модель можна вважати узагальненням понять алгебри й алгебричної системи.

Далі детальніше розглянемо лише один вид алгебричних систем — алгебри. Усі подані нижче означення, властивості й теореми для алгебр можна узагальнити і перенести на випадок алгебричних систем чи моделей.

2.2. Алгебри

Отже, множину M разом із заданою на ній сукупністю операцій $\Omega = \{\varphi_1, \varphi_2, \ldots, \varphi_m, \ldots\}$, тобто систему $A = \langle M, \Omega\rangle$, називають ***алгеброю***; M — її ***носій***, Ω — ***сигнатура*** алгебри A. Кортеж $(n_1, n_2, \ldots, n_m, \ldots)$ арностей операцій сигнатури Ω визначає ***тип*** алгебри A.

Множина $L \subseteq M$ називається ***замкненою*** відносно n-арної операції φ на M, якщо $\mathrm{Pr}_2\varphi_L \subseteq L$, тобто результати операції φ на кортежах із L^n належать L; φ_L — обмеження φ на L.

Якщо множина $L \subseteq M$ замкнена відносно всіх операцій алгебри $A = \langle M, \Omega\rangle$, то систему $B = \langle L, \Omega\rangle$ називають ***підалгеброю*** алгебри A.

Приклад 2.2. 1. Множина R дійсних чисел разом з операціями додавання + і множення $\times$ утворює алгебру $A = \langle R, \{+, \times\}\rangle$. Обидві операції алгебри A бінарні, тому тип цієї алгебри (2, 2). Підалгебрами алгебри A є, зокрема, системи $\langle Q, \{+, \times\}\rangle$, $\langle Z, \{+, \times\}\rangle$ і $\langle N, \{+, \times\}\rangle$.

2. Нехай $N_{(p)} = \{0, 1, 2, \ldots, p-1\}$. Задамо на множині $N_{(p)}$ дві операції $\oplus$ — додавання за модулем p і $\otimes$ — множення за модулем p: $a \oplus b = c$ і $a \otimes b = d$, де c і d — остачі від ділення на p відповідно чисел $a + b$ і $a \times b$, $a, b \in N_{(p)}$. Тоді $A = \langle N_{(p)}, \{\oplus, \otimes\}\rangle$ — алгебра з типом (2, 2). Наприклад, якщо $p = 5$, то $N_{(5)} = \{0, 1, 2, 3, 4\}$ і $2 \oplus 4 = 1$, $1 \oplus 4 = 0$, $2 \otimes 4 = 3$, $1 \otimes 3 = 3$, $0 \otimes 2 = 0$.

Уведені операції $\oplus$ і $\otimes$ часто записують у вигляді $a + b \equiv c \pmod p$, $a \times b \equiv d \pmod p$, а алгебру A називають ***алгеброю остач*** (***лишків***).

3. Нехай $M \subseteq R$ — довільна підмножина дійсних чисел. Тоді система $\langle M, \{\max, \min\}\rangle$ є алгеброю з типом (2, 2). Її підалгебрами є системи $\langle L, \{\max, \min\}\rangle$ для будь-якої підмножини $L \subseteq M$.

4. Нехай $\beta(M)$ — булеан множини M; тоді система $B = \langle\beta(M), \{\cup, \cap, \overline{\ }\}\rangle$ є алгеброю з типом (2, 2, 1), яку називають ***булевою алгеброю множин*** над M. Для унарної операції доповнення $\overline{\ }$ універсальною множиною вважають саму множину M. Підалгебрами алгебри B є системи $B' = \langle\beta(L), \{\cup, \cap, \overline{\ }\}\rangle$ для будь-якої підмножини $L \subseteq M$.

5. Розглянемо множину A^* всіх слів у алфавіті A. *Конкатенацією* (або *множенням*) слів $\alpha_1, \alpha_2 \in A^*$, $\alpha_1 = a_{i1}a_{i2}\ldots a_{ik}$, $\alpha_2 = a_{j1}a_{j2}\ldots a_{jl}$ називається бінарна операція $\circ$, яка парі слів α_1, α_2 ставить у відповідність слово $\alpha_1 \circ \alpha_2$, одержане в результаті приписування праворуч від слова α_1 слова α_2, тобто $\alpha_1 \circ \alpha_2 = a_{i1}a_{i2}\ldots a_{ik}a_{j1}a_{j2}\ldots a_{jl}$. Для порожнього слова e $\alpha \circ e = e \circ \alpha = \alpha$, $\alpha \in A^*$. Система $\langle A^*, \{\circ\}\rangle$ — алгебра типу (2).

6. Нехай M — скінченна множина. Розглянемо сукупність P_M усіх бієктивних відображень множини M на себе, тобто *підстановок* множини M. Тотожну підстановку $i_M = \{(a, a) | a \in M\}$ у цьому разі часто позначають як ε. Остання підстановка займає особливе місце у множині підстановок, і тому її відносять до виділених елементів множини P_M.

Розглянемо на множині підстановок P_M бінарну операцію композиції (добутку) й унарну операцію взяття оберненої підстановки. Добутком $\varphi \circ \psi$ двох підстановок $\varphi, \psi \in P_M$ є підстановка $\eta = \{(a, \psi(\varphi(a))) \mid a \in M\}$, а підстановкою φ^{-1}, оберненою до підстановки $\varphi \in P_M$, є $\{(\varphi(a), a) \mid a \in M\}$. Очевидно, що $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = \varepsilon$ і $\varphi \circ \varepsilon = \varepsilon \circ \varphi = \varphi$ для будь-якої підстановки $\varphi \in P_M$. Система $\Pi = \langle P_M, \{\circ, ^{-1}, \varepsilon\}\rangle$ є ***алгеброю підстановок*** із типом (2, 1, 0).

Оскільки для вивчення алгебр підстановок природа елементів множини M неістотна, то будемо вважати, що $M = \{1, 2, \ldots, n\}$. Підстановку φ з множини P_M записують у вигляді двох рядків. У першому рядку виписують усі елементи множини M, а кожен елемент другого рядка — це образ розміщеного над ним елемента першого рядка при відображенні φ (іноді, вважаючи, що елементи множини M упорядковані, обмежуються записом тільки другого рядка — рядка образів). Наприклад, якщо $n = 5$, то

$$\varepsilon = \begin{pmatrix} 1\,2\,3\,4\,5 \\ 1\,2\,3\,4\,5 \end{pmatrix}, \ \varphi = \begin{pmatrix} 1\,2\,3\,4\,5 \\ 4\,5\,3\,2\,1 \end{pmatrix}, \ \psi = \begin{pmatrix} 1\,2\,3\,4\,5 \\ 2\,4\,3\,1\,5 \end{pmatrix},$$

$$\varphi \circ \psi = \begin{pmatrix} 1\,2\,3\,4\,5 \\ 1\,5\,3\,4\,2 \end{pmatrix}, \ \varphi^{-1} = \begin{pmatrix} 1\,2\,3\,4\,5 \\ 5\,4\,3\,1\,2 \end{pmatrix}.$$

2.1. Задачі і вправи

1. З'ясувати замкненість L відносно операцій сигнатури алгебри $A = \langle N, \{+, \times\}\rangle$, якщо L — це множина:

(а) парних чисел;
(б) непарних чисел;
(в) чисел, кратних k;
(г) чисел виду 2^k, $k \in N$;
(д) чисел виду n^2, $n \in N$;
(е) чисел, більших ніж 1000;
(є) чисел, менших від 1000.

2. Які з множин L попередньої задачі є носіями систем $B = \langle L, \{+, \times\}\rangle$, що утворюють підалгебри алгебри $A = \langle N, \{+, \times\}\rangle$?

3. Написати таблиці додавання та множення алгебри остач $A = \langle N_{(p)}, \{\oplus, \otimes\}\rangle$ для таких значень p:

(а) $p = 3$; (б) $p = 4$; (в) $p = 7$; (г) $p = 2$.

4. Чи є операції $\oplus$ і $\otimes$ алгебри остач:

(а) асоціативними;
(б) комутативними;
(в) дистрибутивними одна відносно одної?

5. З'ясувати, чи є операція конкатенації слів:

(а) асоціативною; (б) комутативною?

6. Довести, що операція композиції алгебри підстановок Π асоціативна та некомутативна.

7. Скільки елементів містить носій P_M алгебри підстановок Π, якщо $|M| = n$?

8. Для $M = \{1, 2, \ldots, n\}$ розглянемо множину P'_M підстановок f таких, що $f(i) \equiv i + k \pmod n$, $i = 1, 2, \ldots, n$, $k = 0, 1, 2, \ldots$ (тобто нижній рядок запису підстановки f є результатом зсуву чисел $1, 2, \ldots, n$ на k позицій праворуч). Визначити:

(а) кількість елементів у множині P'_M;

(б) чи утворює $\Pi' = \langle P'_M, \{\circ, ^{-1}, \varepsilon\}\rangle$ підалгебру алгебри підстановок Π,

(в) чи є операція композиції комутативною в Π';

(г) чи існує для будь-якої підстановки $f \in P'_M$ таке додатне число m, що $f^{(m)} = \varepsilon$.

9. Для будь-яких цілих чисел n і m ($m \in N$) позначимо через $r(n, m)$ остачу від ділення n на m. Довести, що для $m \geq 2$ і числа k, взаємно простого з m, відповідність $f(n) = r(kn, m)$ є підстановкою множини $M = \{0, 1, 2, \ldots, m-1\}$.

10. Довести, що в алгебрі $\Pi = \langle P_M, \{\circ, ^{-1}, \varepsilon\}\rangle$ для довільних підстановок $f, g \in P_M$

(а) існує таке додатне число m, що $f^{(m)} = \varepsilon$;

(б) існує така підстановка h, що $f \circ h = g$;

(в) існує така підстановка w, що $w \circ f = g$.

11. Означимо на множині M за допомогою підстановки $f \in P_M$ відношення R так: $a\,R\,b$ тоді й тільки тоді, коли для деякого невід'ємного цілого m виконується $f^{(m)}(a) = b$. Довести, що R — еквівалентність на M.

2.3. Гомоморфізм та ізоморфізм алгебр

Для зручності припустімо, що елементи типів усіх розглядуваних нижче алгебр певним чином упорядковані.

Алгебри називають ***однотипними***, якщо вони мають однакові типи.

Нехай $A_1 = \langle M_1, \Omega_1\rangle$ і $A_2 = \langle M_2, \Omega_2\rangle$ — однотипні алгебри; $\Omega_1 = \{\varphi_1, \varphi_2, \ldots, \varphi_m, \ldots\}$, $\Omega_2 = \{\psi_1, \psi_2, \ldots, \psi_m, \ldots\}$, l_m — арність операцій φ_m і ψ_m (за умовою арності операцій φ_m і ψ_m однакові).

Відображення γ: $M_1 \to M_2$ називається ***гомоморфізмом***, або ***гомоморфним відображенням***, алгебри A_1 ***в*** алгебру A_2, якщо

$$\gamma\big(\varphi_i(a_{j1}, a_{j2}, \ldots, a_{jl_i})\big) = \psi_i\big(\gamma(a_{j1}), \gamma(a_{j2}), \ldots, \gamma(a_{jl_i})\big) \qquad (2.1)$$

для всіх $i = 1, 2, \ldots, m, \ldots$; $a_{jk} \in M$, $k = 1, 2, \ldots, l_i$.

Змістовно наведене означення можна інтерпретувати так: відображення γ носія M_1 у носій M_2 двох однотипних алгебр A_1 і A_2 є гомоморфізмом алгебри A_1 в алгебру A_2, якщо образ результата будь-якої операції $\varphi_i \in \Omega_1$ із довільними аргументами $a_{j1}, a_{j2}, \ldots, a_{jl_i}$ збігається з результатом відповідної операції $\psi_i \in \Omega_2$ з аргументами, які є образами елементів $a_{j1}, a_{j2}, \ldots, a_{jl_i}$ при відображенні γ.

Якщо для однотипних алгебр $A_1 = \langle M_1, \Omega_1\rangle$ і $A_2 = \langle M_2, \Omega_2\rangle$ існує взаємно однозначне відображення γ: $M_1 \to M_2$, для якого виконуються умови (2.1), то γ називають ***ізоморфізмом***, або ***ізоморфним відображенням***, алгебри A_1 ***на*** алгебру A_2. При цьому алгебри A_1 і A_2 називають ***ізоморфними***. Якщо $M_1 = M_2$, то ізоморфізм називається ізоморфізмом ***на*** себе, або ***автоморфізмом***; якщо $M_2 \subseteq M_1$, то ізоморфізм називається ***ізоморфізмом у*** себе.

Приклад 2.3. 1. Нехай Z — множина цілих чисел, а $Z_{(k)}$ — множина всіх цілих чисел, кратних k. Відображення γ: $Z \to Z_{(k)}$ таке, що $\gamma(n) = kn$, $n \in Z$, є ізоморфізмом алгебри $\langle Z, \{+\}\rangle$ на алгебру $\langle Z_{(k)}, \{+\}\rangle$. Умова (2.1) набирає вигляду $\gamma(n + m) = k(n + m) = kn + km = \gamma(n) + \gamma(m)$, $n, m \in Z$. Оскільки $Z_{(k)} \subseteq Z$, то γ — ізоморфізм у себе.

Відображення γ: $Z \to Z$ за правилом $\gamma(n) = -n$, $n \in Z$, — автоморфізм алгебри $\langle Z, \{+\}\rangle$. Умова (2.1) тут має вигляд $\gamma(n + m) = -(n + m) = (-n) + (-m) = \gamma(n) + \gamma(m)$. Водночас для алгебри $\langle Z, \{\times\}\rangle$ γ — не автоморфізм, бо $-(n \times m) \neq (-n) \times (-m)$, $n, m \in Z$.

2. Розглянемо алгебри $\langle N, \{+, \times\}\rangle$ та $\langle N_{(p)}, \{\oplus, \otimes\}\rangle$ (див. приклад 2.2.2). Задамо відображення γ: $N \to N_{(p)}$ так: $\gamma(n) = r$, де r — остача від ділення n на p. Неважко переконатися, що виконуються співвідношення $\gamma(n + m) = \gamma(n) \oplus \gamma(m)$ і $\gamma(n \times m) = \gamma(n) \otimes \gamma(m)$, тобто γ — гомоморфізм алгебри $\langle N, \{+, \times\}\rangle$ в алгебру $\langle N_{(p)}, \{\oplus, \otimes\}\rangle$.

Цей приклад показує, що можливий гомоморфізм алгебри з нескінченним носієм N в алгебру зі скінченним носієм $N_{(p)}$.

Побудований гомоморфізм γ ініціює відношення R_p на множині N: nR_pm тоді й тільки тоді, коли $\gamma(n) = \gamma(m)$; R_p — відношення конгруентності за модулем p, отже, є відношення еквівалентності на N. За цим відношенням множина N розбивається на p підмножин $L_k = \{n \mid n \in N \text{ і } \gamma(n) = k\}$, $k = 0, 1, 2, \ldots, p - 1$ (див. приклад 1.16.2).

3. Нехай R — множина дійсних чисел, а R^+ — множина додатних дійсних чисел. Ізоморфізм алгебри $\langle R, \{+\}\rangle$ на алгебру $\langle R^+, \{\times\}\rangle$ можна задати за допомогою відображення $\gamma(x) = a^x$, $x \in R$ ($a > 0$, $a \neq 1$). Умова (2.1) набирає вигляду $\gamma(x + y) = a^{x+y} = a^x \times a^y = \gamma(x) \times \gamma(y)$, $x, y \in R$.

І навпаки, ізоморфізм $\langle R^+, \{\times\}\rangle$ на $\langle R, \{+\}\rangle$ можна задати відображенням $\gamma(x) = \log_a x$, $x \in R^+$ ($a > 0$, $a \neq 1$). Обидва ці ізоморфізми успішно використовували протягом сторіч, працюючи з так званою *логарифмічною лінійкою* й іншими стародавніми механічними пристроями для полегшення та спрощення обчислень за допомогою зведення операції множення чисел до операції додавання.

4. Будь-які булеві алгебри множин $\langle \beta(M), \{\cup, \cap, \bar{\ }\}\rangle$ і $\langle \beta(M'), \{\cup, \cap, \bar{\ }\}\rangle$ для рівнопотужних множин M і M' ізоморфні між собою. Ізоморфним відображенням може бути будь-яке взаємно однозначне відображення γ з M на M'; операції обох алгебр однакові.

Неважко переконатися, що ізоморфізм — це відношення еквівалентності на множині алгебр.

Розглянемо алгебру $A_1 = \langle B^n, \{\vee, \wedge, \neg\}\rangle$. Тут $B = \{0, 1\}$ — двійковий алфавіт і B^n — множина двійкових кортежів довжини n, а бінарні операції $\vee$ та $\wedge$ й унарну операцію $\neg$ для довільних двійкових кортежів $\alpha = (b_1, b_2, \ldots, b_n)$ і $\beta = (d_1, d_2, \ldots, d_n)$ означимо так:

$$\alpha \vee \beta = (b_1 \vee d_1, b_2 \vee d_2, \ldots, b_n \vee d_n),$$
$$\alpha \wedge \beta = (b_1 \wedge d_1, b_2 \wedge d_2, \ldots, b_n \wedge d_n),$$
$$\neg\alpha = (\neg b_1, \neg b_2, \ldots, \neg b_n),$$

тобто вони виконуються покомпонентно, $b_i, d_i \in B$, $i = 1, 2, \ldots, n$. Означимо операції *диз'юнкції* $\vee$ та *кон'юнкції* $\wedge$ й операцію *заперечення* $\neg$ на множині B. Вважаємо, що $b \vee d = 0$ тоді й тільки тоді, коли $b = d = 0$, в іншому випадку — $b \vee d = 1$. Покладемо $b \wedge d = 1$ тоді й тільки тоді, коли $b = d = 1$, в іншому разі $b \wedge d = 0$. Унарну операцію $\neg b$ означено рівностями $\neg 0 = 1$ і $\neg 1 = 0$.

Нехай $M = \{a_1, a_2, \ldots, a_n\}$ — скінченна множина, $A_2 = \langle \beta(M), \{\cup, \cap, \bar{\ }\}\rangle$ — булева алгебра множин над M. Існує взаємно однозначна відповідність γ між множинами $\beta(M)$ і B^n (див. доведення теореми 1.1). Будь-якій підмножині $L \subseteq M$ при бієктивному відображенні γ відповідає двійковий кортеж $\alpha = (c_1, c_2, \ldots, c_n)$ такий, що $c_i = 1$, коли $a_i \in L$ і $c_i = 0$, коли $a_i \notin L$.

Справедлива така теорема.

Теорема 2.1. Якщо $M = \{a_1, a_2, \ldots, a_n\}$ — скінченна множина, то алгебра $\langle \beta(M), \{\cup, \cap, \bar{\ }\}\rangle$ ізоморфна алгебрі $\langle B^n, \{\vee, \wedge, \neg\}\rangle$.

Доведення. Доведемо, що означене вище відображення γ: $\beta(M) \to B^n$ задовольняє умови (2.1) і, отже, є ізоморфізмом.

Нехай $M_1, M_2 \subseteq M$ — довільні підмножини множини M, $\gamma(M_1) = \alpha = (b_1, b_2, \ldots, b_n)$ і $\gamma(M_2) = \beta = (d_1, d_2, \ldots, d_n)$. Тоді $\gamma(M_1 \cup M_2) = \sigma = (c_1, c_2, \ldots, c_n)$, де $c_i = 0$ тоді й тільки тоді, коли $a_i \notin M_1 \cup M_2$. Останнє означає, що $a_i \notin M_1$ і $a_i \notin M_2$, тобто $b_i = d_i = 0$. Отже, $\sigma = \alpha \vee \beta$.

Аналогічно, якщо $\gamma(M_1 \cap M_2) = \tau = (h_1, h_2, \ldots, h_n)$, то $h_i = 1$ тоді й лише тоді, коли $a_i \in M_1 \cap M_2$. Тоді $a_i \in M_1$, $a_i \in M_2$ і $b_i = d_i = 1$. Тому $\tau = \alpha \wedge \beta$.

Нарешті, нехай $\gamma(\overline{M}_1) = \delta = (l_1, l_2, \ldots, l_n)$. Тоді $l_i = 1$, якщо $a_i \in \overline{M}_1$, тобто $a_i \notin M_1$ і $b_i = 0$; навпаки, $l_i = 0$, якщо $a_i \notin \overline{M}_1$, тобто $a_i \in M_1$ і $b_i = 1$. Це означає, що $\delta = \neg\alpha$. Отже, виконуються всі потрібні співвідношення:

$$\gamma(M_1 \cup M_2) = \gamma(M_1) \vee \gamma(M_2),$$
$$\gamma(M_1 \cap M_2) = \gamma(M_1) \wedge \gamma(M_2),$$
$$\gamma(\overline{M}_1) = \neg\gamma(M_1).$$

Теорема 2.1 має особливе значення, зокрема, для програмування на ЕОМ. Якщо у програмі потрібно реалізувати операції об'єднання, перетину чи доповнення для якихось скінченних множин, то ці операції можна промоделювати за допомогою відповідних покомпонентних операцій над певними двійковими кортежами. Зображення двійкових кортежів і реалізація порозрядних операцій над ними в сучасних ЕОМ виконуються дуже просто й ефективно.

Поняття ізоморфізму — одне з найважливіших у математиці. З умов (2.1) випливає, що будь-яке тотожне співвідношення для елементів алгебри A має відповідне йому тотожне співвідношення у будь-якій ізоморфній їй алгебрі A'. Ця властивість ізоморфних алгебр дає змогу, одержавши певний результат (тотожність, властивість, теорему тощо) стосовно елементів алгебри A, автоматично поширювати його на елементи всіх ізоморфних алгебр. На цій підставі кажуть, що за внутрішнім "устроєм", або внутрішньою структурою, усі ізоморфні алгебри однакові, і достатньо детально вивчити лише одного з "представників" таких алгебр. Завдяки встановленню ізоморфізму між об'єктами, на перший погляд, здавалося б, абсолютно різної природи з далеких між собою розділів математики було одержано багато красивих і глибоких результатів. Поширений у математиці вираз “із точністю до ізоморфізму” означає, що розглядають і досліджують ті властивості математичних об’єктів, які зберігаються при встановленні ізоморфізму, тобто є спільними для всіх ізоморфних об’єктів.

2.2. Задачі і вправи

1. Довести, що існує безліч підалгебр $A_2 = \langle M, \{+, \times\}\rangle$, ізоморфних алгебрі $A_1 = \langle Z, \{+, \times\}\rangle$, $M \subseteq Z$.

2. Довести ізоморфізм алгебр A_1 і A_2:

(а) $A_1 = \langle R^+, \{+\}\rangle$, $A_2 = \langle R^-, \{+\}\rangle$;

(б) $A_1 = \langle M_1, \{\times\}\rangle$ і $A_2 = \langle M_2, \{\times\}\rangle$, де $M_1 = \{p \mid p\in Q,\ 0<p<1\}$ і $M_2 = \{p \mid p\in Q,\ p>1\}$.

3. Довести ізоморфізм алгебр $A_1 = \langle M_1, \Omega_1\rangle$ і $A_2 = \langle M_2, \Omega_2\rangle$, якщо

(а) M_1 — множина матриць виду $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, $a, b\in R$, Ω_1 включає дві операції: додавання і множення матриць; M_2 — множина комплексних чисел $a + bi$, а $\Omega_2 = \{+, \times\}$;

(б) M_1 — множина матриць виду $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, $a\in R$, Ω_1 включає дві бінарні операції додавання і множення матриць і унарну операцію обчислення оберненої матриці; $M_2 = R$, $\Omega_2 = \{+, \times, ^{-1}\}$;

(в) M_1 — множина матриць виду $\begin{pmatrix} a & b \\ 2b & a \end{pmatrix}$, $a, b\in Q$, Ω_1 містить операції додавання і множення матриць; M_2 — множина чисел виду $a + b\sqrt{2}$, $a, b\in Q$, $\Omega_2 = \{+, \times\}$;

(г) M_1 — множина матриць виду $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, $a, b\in R$, Ω_1 включає операцію додавання матриць і операцію множення матриці на дійсне число; M_2 — множина векторів (a, b), $a, b\in R$, Ω_2 містить операцію додавання векторів і операцію множення вектора на дійсне число;

(д) M_1 — множина векторів (a, b), $a, b\in R$, Ω_1 включає операцію додавання векторів $(a, b) + (c, d) = (a + c, b + d)$ і операцію множення $(a, b)\otimes(c, d) = (ac + bd, ad + bc)$; M_2 — множина матриць виду $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$, $a, b\in R$, Ω_2 складається з операцій додавання та множення матриць.

4. Довести, що алгебри A_1 і A_2 не є ізоморфними

(а) $A_1 = \langle R, \{+, \times\}\rangle$, $A_2 = \langle Q, \{+, \times\}\rangle$;

(б) $A_1 = \langle R, \{+, \times\}\rangle$, $A_2 = \langle C, \{+, \times\}\rangle$, C — множина комплексних чисел;

(в) $A_1 = \langle Z, \{+, \times\}\rangle$, $A_2 = \langle N, \{+, \times\}\rangle$;

(г) $A_1 = \langle M_1, \{+, \times\}\rangle$, $A_2 = \langle M_2, \{+, \times\}\rangle$, де $M_1 = \{a + b\sqrt{2} \mid a, b\in Q\}$ і $M_2 = \{a + b\sqrt{3} \mid a, b\in Q\}$;

(д) $A_1 = \langle R, \{+\}\rangle$, $A_2 = \langle R, \{\times\}\rangle$;

(е) $A_1 = \langle Q, \{+\}\rangle$, $A_2 = \langle Q, \{\times\}\rangle$;

(є) $A_1 = \langle N, \{+\}\rangle$, $A_2 = \langle N, \{\times\}\rangle$.

5. Побудувати нетривіальний (нетотожний) автоморфізм алгебри $A = \langle M, \{+, \times\}\rangle$, де $M = \{a + b\sqrt{2} \mid a, b\in Q\}$.

6. Довести, що гомоморфізм γ алгебри $A_1 = \langle M_1, \Omega_1\rangle$ в алгебру $A_2 = \langle M_2, \Omega_2\rangle$ буде ізоморфізмом тоді й тільки тоді, коли існує такий гомоморфізм φ алгебри A_2 в алгебру A_1, що $\gamma \circ \varphi$ та $\varphi \circ \gamma$ — тотожні перетворення відповідно но-сіїв M_1 і M_2.

2.4. Фактор-алгебра. Теорема про гомоморфізми

Відношення еквівалентності R на множині M називається ***конгруенцією*** в алгебрі $A = \langle M, \Omega\rangle$, якщо для будь-якої n-арної операції $\varphi\in\Omega$ та будь-яких елементів a_i, $b_i \in M$ таких, що $a_i R\, b_i$, $i = 1, 2, \ldots, n$, виконується співвідношення $\varphi(a_1, a_2, \ldots, a_n)\ R\ \varphi(b_1, b_2, \ldots, b_n)$.

Інакше кажучи, якщо $B_1, B_2, \ldots, B_n$ — класи розбиття множини M за відношенням еквівалентності R (елементи фактор-множини M/R), то клас C, якому належить елемент $\varphi(a_1, a_2, \ldots, a_n)$, $a_i\in B_i$, $i = 1, 2, \ldots, n$, не залежить від вибору елементів a_i в класах B_i, $i = 1, 2, \ldots, n$. Це дає змогу для кожної n-арної операції $\varphi\in\Omega$ на множині M означити відповідну n-арну операцію φ' на фактор-множині M/R рівністю

$$\varphi'(B_1, B_2, \ldots, B_n) = C. \tag{2.2}$$

У такий спосіб можна утворити нову алгебру $A/R = \langle M/R, \Omega'\rangle$, $\Omega' = \{\varphi' \mid \varphi\in\Omega\}$, однотипну з алгеброю A (найчастіше для операцій з Ω' зберігають ті самі позначення, що і в Ω). Алгебра A/R називається ***фактор-алгеброю*** алгебри A за конгруенцією R.

Розглянемо канонічне відображення γ_R: $M \to M/R$. Із принципу означення операцій (2.2) у фактор-алгебрі A/R і співвідношень (2.1) випливає, що γ_R є гомоморфізмом алгебри A на фактор-алгебру A/R. Цей гомоморфізм називають ***канонічним***, або ***природним***, ***гомоморфізмом*** A на A/R.

Приклад 2.4. Для деякого фіксованого натурального числа p розглянемо відношення R на множині N: mRn тоді й тільки тоді, коли $m \equiv n \pmod p$. Неважко переконатися, що R — конгруенція в алгебрі $A = \langle N, \{+, \times\}\rangle$. Класи еквівалентності за відношенням R — це $[0]_R$, $[1]_R$, ..., $[p-1]_R$, де $[k]_R$ — множина всіх натуральних чисел, які при діленні на p дають остачу k. Тут $\Omega' = \{\oplus, \otimes\}$, де операції $\oplus$ та $\otimes$ означено співвідношеннями $[k]_R \oplus [l]_R = [k+l]_R$, $[k]_R \otimes [l]_R = [k \times l]_R$. Отже, алгебра $A/R = \langle N/R, \{\oplus, \otimes\}\rangle$, де $N/R = \{[0]_R, [1]_R, \ldots, [p-1]_R\}$ є фактор-алгеброю алгебри A за конгруенцією R, а відображення $\gamma_R(k) = [k]_R$ — канонічним гомоморфізмом A на A/R. Алгебра A/R ізоморфна алгебрі остач $\langle N_{(p)}, \{\oplus, \otimes\}\rangle$.

Оберненим до наведеного вище твердження є *теорема про гомоморфізми*, яка описує загальні властивості всіх гомоморфізмів певної алгебри.

Теорема 2.2. Якщо $A_1 = \langle M_1, \Omega_1\rangle$ і $A_2 = \langle M_2, \Omega_2\rangle$ — однотипні алгебри і γ — гомоморфізм алгебри A_1 в алгебру A_2, то в алгебрі A_1 існує така конгруенція R, що алгебра A_2 та фактор-алгебра A_1/R ізоморфні між собою. Крім того, існує таке ізоморфне відображення δ алгебри A_2 на фактор-алгебру A_1/R, що відображення $\gamma \circ \delta$ збігається з канонічним гомоморфізмом A_1 на A_1/R.

Доведення. Розглянемо на множині M_1 таке відношення R: $a\,R\,b$ тоді й тільки тоді, коли $\gamma(a) = \gamma(b)$. Відношення R є еквівалентністю на M_1. Більше того, R — конгруенція в алгебрі A_1. Справді, для будь-якої n-арної операції $\varphi \in \Omega_1$ і кортежів аргументів $(a_1, a_1, \ldots, a_n)$ і $(b_1, b_2, \ldots, b_n)$ таких, що $a_i\,R\,b_i$ або $\gamma(a_i) = \gamma(b_i)$, $i = 1, 2, \ldots, n$, маємо $\gamma(\varphi(a_1, a_2, \ldots, a_n)) = = \psi(\gamma(a_1), \gamma(a_2), \ldots, \gamma(a_n)) = \psi(\gamma(b_1), \gamma(b_2), \ldots, \gamma(b_n)) = \gamma(\varphi(b_1, b_2, \ldots, b_n))$, отже $\varphi(a_1, a_2, \ldots, a_n)\,R\,\varphi(b_1, b_2, \ldots, b_n)$.

Класами еквівалентності за відношенням R у множині M_1 є множини $\gamma^{-1}(a)$ для кожного $a \in \mathrm{Pr}_2\gamma \subseteq M_2$. Отже, ставлячи у відповідність кожному елементу $a \in \mathrm{Pr}_2\gamma$ його прообраз $\gamma^{-1}(a)$ при гомоморфізмі γ, одержуємо взаємно однозначне відображення δ множини M_2 на фактор-множину M_1/R. Доведемо, що δ — ізоморфізм для відповідних алгебр A_2 та A_1/R, тобто доведемо, що для δ виконуються співвідношення (2.1).

Нехай $\varphi \in \Omega_1$ — довільна n-арна операція алгебри A_1, а $\varphi' \in \Omega_1'$ і $\psi \in \Omega_2$ — n-арні операції фактор-алгебри A_1/R і алгебри A_2, які відповідають операції φ. Для довільних елементів $b_1, b_2, \ldots, b_n \in \mathrm{Pr}_2\gamma \subseteq M_2$ існують такі $a_1, a_2, \ldots, a_n \in M_1$, що $a_i \in \gamma^{-1}(b_i)$, тобто $\gamma(a_i) = b_i$, $i = 1, 2, \ldots, n$. Тоді $\delta(\psi(b_1, b_2, \ldots, b_n)) = \delta(\psi(\gamma(a_1), \gamma(a_2), \ldots, \gamma(a_n))) = \delta(\gamma(\varphi(a_1, a_2, \ldots, a_n)))$.

Оскільки $\varphi(a_1, a_2, \ldots, a_n) \in \varphi'(\gamma^{-1}(b_1), \gamma^{-1}(b_2), \ldots, \gamma^{-1}(b_n))$, то $\delta(\psi(\gamma(a_1), \gamma(a_2), \ldots, \gamma(a_n))) = \varphi'(\gamma^{-1}(b_1), \gamma^{-1}(b_2), \ldots, \gamma^{-1}(b_n))$.

Отже, $\delta(\psi(b_1, b_2, \ldots, b_n)) = \varphi'(\gamma^{-1}(b_1), \gamma^{-1}(b_2), \ldots, \gamma^{-1}(b_n)) = \varphi'(\delta(b_1), \delta(b_2), \ldots, \delta(b_n))$ і δ — ізоморфізм алгебр A_2 та A_1/R. Нарешті, якщо a — довільний елемент множини M_1 і $\gamma(a) = b$, $\delta(b) = \gamma^{-1}(b)$, то $a \in \gamma^{-1}(b)$, тобто $\gamma \circ \delta$ — канонічний гомоморфізм алгебри A_1 на фактор-алгебру A_1/R.

2.3. Задачі і вправи

1. Довести, що відношення R з прикладу 2.4 — конгруенція.

2. Довести ізоморфізм фактор-алгебри A/R з прикладу 2.4 та відповідної алгебри остач.

3. Довести, що відношення рівності i_M є конгруенцією в будь-якій алгебрі $A = \langle M, \Omega\rangle$.

4. На множині A^* всіх слів у алфавіті A розглянемо відношення L: $w_1\,L\,w_2$ тоді й тільки тоді, коли довжини слів w_1 і w_2 збігаються. Довести, що L — конгруенція в алгебрі $W = \langle A^*, \{\circ\}\rangle$, і побудувати фактор-алгебру W/L.

5. Означимо на множині A^* всіх слів у алфавіті A відношення F: $w_1\,F\,w_2$ тоді й тільки тоді, коли перші символи слів w_1 і w_2 збігаються. Довести, що F — конгруенція в алгебрі $W = \langle A^*, \{\circ\}\rangle$ і побудувати відповідну фактор-алгебру W/F.

6. В n-вимірному векторному просторі R^n з операцією $+$ додавання векторів і операцією $\times$ множення вектора на дійсне число означимо відношення P: $(a_1, a_2, \ldots, a_n)\,P\,(b_1, b_2, \ldots, b_n)$ тоді й тільки тоді, коли $a_1 = b_1$. Довести, що P — конгруенція в алгебрі $A = \langle R^n, \{+, \times\}\rangle$, і побудувати відповідну фактор-алгебру A/P.

7. Довести, що фактор-алгебра A/P з попереднього прикладу ізоморфна алгебрі $\langle R, \{+, \times\}\rangle$.

8. В n-вимірному нормованому векторному просторі для елементів $a, b \in R^n$ означимо відношення D: $a\,D\,b$ тоді й тільки тоді, коли $\|a\| = \|b\|$, де $\|a\|$ — норма вектора a. Чи є відношення D

(а) еквівалентністю на R^n;

(б) конгруенцією в алгебрі $A = \langle R^n, \{+, \times\}\rangle$?

9. Розглянемо алгебру $A = \langle M, \{+, -, \times, :\}\rangle$, в якій $M = Z \times N$, а операції сигнатури означено так: $(a, b) \pm (c, d) = (ad \pm bc, bd)$, $(a, b) \times (c, d) = (ac, bd)$,

$(a, b) : (c, d) = (ad, bc)$, якщо $c \neq 0$. Означимо на M відношення R: $(a, b)\ R\ (c, d)$ тоді й тільки тоді, коли $ad = bc$. Довести, що R — конгруенція в алгебрі A, і побудувати відповідну фактор-алгебру A/R.

10. Довести, що фактор-алгебра A/R з останнього прикладу ізоморфна алгебрі $B = \langle Q, \{+, -, \times, /\}\rangle$.

2.5. Системи твірних

Розглянемо алгебру $A = \langle M, \Omega\rangle$, $\Omega = \{\varphi_1, \varphi_2, \ldots, \varphi_m, \ldots\}$. На операціях сигнатури Ω можна означити ***оператор суперпозиції*** S, який дає змогу з основних операцій φ_i, $i = 1, 2, \ldots$, одержувати нові операції на множині M.

Означимо поняття суперпозиції функцій індуктивно:

а) вважаємо суперпозиціями всі функції (операції) сигнатури Ω;

б) нехай задано сукупність $S_1, S_2, \ldots, S_n$, де S_i — або деяка змінна, що набуває значення у множині M, або вже побудована раніше суперпозиція, $i = 1, 2, \ldots, n$, й f — n-арна операція сигнатури Ω. Тоді суперпозицією $S(S_1, S_2, \ldots, S_n, f)$ вважаємо означену на M функцію $f(S_1, S_2, \ldots, S_n)$.

Вираз, що описує суперпозицію та містить функціональні символи чи знаки операцій, круглі дужки й символи аргументів, називається ***формулою*** в алгебрі A.

Якщо f деяка n-арна функція на множині M, то за допомогою операції суперпозиції S можна здійснити перейменування аргументів функції f, зокрема ототожнювання й переставлення цих аргументів. При цьому одержуємо функції, взагалі кажучи, різних типів і відповідно різні формули.

Приклад 2.5. 1. Суперпозицією $S(S_1, S_2, f)$ функцій $S_1(x_1, x_2) = x_1 \times x_2$, $S_2(x_3) = x_3^2$ і $f(x_1, x_2) = x_1 + x_2$ є функція $\varphi(x_1, x_2, x_3) = f(S_1(x_1, x_2), S_2(x_3)) = x_1 \times x_2 + x_3^2$ типу $M^3 \to M$. Якщо ж $S_1(x_1, x_2) = x_1 \times x_2$, $S_2(x_1) = x_1^2$ і $f(x_1, x_2) = x_1 + x_2$, то суперпозиція $S(S_1, S_2, f)$ — це функція $\psi(x_1, x_2) = x_1 \times x_2 + x_1^2$ типу $M^2 \to M$.

2. Нехай $f(x_1, x_2, x_3) = x_1 + x_2 - x_3$. Якщо $S_1(x_2) = x_2$, $S_2(x_3) = x_3$ і $S_3(x_1) = x_1$, то результат суперпозиції $S(S_1, S_2, S_3, f)$ — функція $\varphi_1(x_1, x_2, x_3) = x_2 + x_3 - x_1$ типу $M^3 \to M$. Для $S_1(x_1) = x_1$, $S_2(x_2) = x_2$ і $S_3(x_2) = x_2$ суперпозиція $S(S_1, S_2, S_3, f)$ дає функцію $\varphi_2(x_1) = x_1$ типу $M \to M$.

Функції, утворені за допомогою оператора суперпозиції з основних операцій алгебри $A = \langle M, \Omega\rangle$, називають ***похідними операціями*** алгебри A. Множину всіх похідних операцій алгебри $A = \langle M, \Omega\rangle$ назвемо ***замиканням*** множини Ω за оператором суперпозиції S і позначатимемо $[\Omega]_S$. Алгебру $A_S = \langle M, [\Omega]_S\rangle$ будемо називати ***похідною*** від алгебри $A = \langle M, \Omega\rangle$.

Нехай $M_1 \subseteq M$. Множина $[M_1]$ називається ***замиканням*** множини M_1 в алгебрі $A = \langle M, \Omega\rangle$, якщо $[M_1]$ складається з усіх елементів, які є результатами операцій сигнатури $[\Omega]_S$ похідної алгебри A_S на аргументах з множини M_1. Вважаємо, що $M_1 \subseteq [M_1]$.

Множина M_1 називається ***замкненою***, якщо $[M_1] = M_1$.

Сукупність елементів $T \subseteq M$ називається ***системою твірних*** (або *повною системою елементів*) алгебри $A = \langle M, \Omega\rangle$, якщо $[T] = M$.

Для будь-якої алгебри $A = \langle M, \Omega\rangle$ існує *тривіальна* система твірних — сам носій M. Для пошуку нетривіальних систем твірних алгебри A важливо мати ефективний критерій, який давав би змогу для певної сукупності елементів з M з'ясовувати, чи є вона системою твірних алгебри A. Сформульована проблема називається ***проблемою повноти*** для даної алгебри, а відповідний критерій — ***критерієм повноти***.

Система твірних B алгебри $A = \langle M, \Omega\rangle$ називається ***базисом*** алгебри A, якщо $[B \backslash \{b\}] \neq M$ для кожного елемента $b \in B$. Інакше кажучи, сукупність елементів B є базисом алгебри $A = \langle M, \Omega\rangle$, якщо $[B] = M$ і для кожного елемента $b \in B$ виконується $b \notin [B \backslash \{b\}]$.

Особливий інтерес при вивченні алгебричних систем становлять алгебри, для яких існують скінченні системи твірних і скінченні базиси.

Алгебра A називається ***скінченно-породжуваною***, якщо вона має скінченну систему твірних, в іншому разі алгебра A *нескінченно-породжувана*.

Наприклад, при проектуванні та побудові різноманітних автоматів (зокрема, електронних обчислювальних машин) використовують набір заданих стандартних блоків, або так званих елементарних автоматів. Аналогічно при написанні програм для ЕОМ на якійсь алгоритмічній мові користуються певним набором основних операторів. Множини елементарних автоматів і основних операторів разом із сукупностями операцій, які задають правила конструювання з елементарних автоматів і основних операторів довільних автоматів і відповідно довільних програм, утворюють дві алгебри. Природно виникає питання: чи існують нетривіальні (бажано скінченні) системи твірних для таких алгебр? У разі позитивної відповіді на ці питання постає проблема пошуку відповідних систем твірних чи базисів [7; 9].

Приклад 2.6. 1. Будь-яка множина натуральних чисел T, що містить число 1, є системою твірних алгебри $\langle N, \{+\}\rangle$. Ця алгебра має єдиний базис $B = \{1\}$, отже вона скінченно-породжувана.

Аналогічно, системою твірних алгебри $\langle Z, \{+\}\rangle$, носієм якої є множина цілих чисел Z, може бути, наприклад, множина $B = \{-1, 1\}$. Остання множина — скінченний базис алгебри $\langle Z, \{+\}\rangle$.

2. Алгебра $\langle N, \{\times\}\rangle$ не має скінченної системи твірних і є нескінченно-породжуваною. Справді, припустімо, що існує скінченна множина T натуральних чисел таких, що за допомогою операції множення з них можна одержати всі натуральні числа. Однак для будь-яких $p, q \in T$ число $p \times q$ не просте, отже, множина простих чисел $P \subseteq N$ не міститься у множині $[T]$, тобто $[T] \neq N$. Саме множина всіх простих чисел P утворює нескінченний базис алгебри $\langle N, \{\times\}\rangle$. Це випливає з того, що будь-яке натуральне число n можна подати у вигляді $n = p_1^{k_1} \times p_2^{k_2} \times \ldots \times p_s^{k_s}$, де $p_1, p_2, \ldots, p_s \in P$, а $p_i^{k_i}$ — скорочення запису виразу $p_i \times p_i \times \ldots \times p_i$ (k_i разів). Крім того, жодне просте число не можна подати у вигляді добутку інших чисел, тому для кожного числа $p \in P$ маємо $p \notin [P \setminus \{p\}]$.

3. Системою твірних алгебри $\langle A^*, \{^\circ\}\rangle$ (див. приклад 2.2.5) є будь-яка множина T така, що $E \subseteq T$, де $E = \{e\} \cup \{a_i \mid a_i \in A, i = 1, 2, \ldots, n\}$ — множина, яка складається з порожнього слова e й усіх односимвольних слів у алфавіті A. Множина E є скінченним базисом алгебри $\langle A^*, \{^\circ\}\rangle$.

4. Базисом булевої алгебри множин $A = \langle \beta(M), \{\cup, \cap, ^{-}\}\rangle$ над довільною множиною M є $M' = \{\{a\} \mid a \in M\}$. Отже, булева алгебра множин скінченно-породжувана тільки для скінченних множин M.

5. Нарешті, наведемо приклад алгебри, що має безліч систем твірних, але не має жодного базису. Розглянемо алгебру $A = \langle N, \{\,'\,\}\rangle$, носій якої — множина натуральних чисел, а єдина операція сигнатури — унарна операція $'$ така, що $1' = 1$, а для всіх інших чисел $n \in N \setminus \{1\}$ виконується $n' = n - 1$. Будь-яка нескінченна підмножина $T \subseteq N$ є системою твірних цієї алгебри. Водночас, у кожній системі твірних T є безліч таких елементів $m \in T$, що $m \in [T \setminus \{m\}]$. Отже, алгебра A не має базису.

2.4. Задачі і вправи

1. Для алгебри $A = \langle M, \Omega\rangle$ і множин $M_1 \subseteq M$ і $M_2 \subseteq M$ обґрунтувати такі властивості замикання:

(а) $[\varnothing] = \varnothing$;

(б) $[[M_1]] = [M_1]$;

(в) якщо $M_1 \subseteq M_2$, то $[M_1] \subseteq [M_2]$;

(г) $[M_1 \cap M_2] \subseteq [M_1] \cap [M_2]$;

(д) $[M_1] \cup [M_2] \subseteq [M_1 \cup M_2]$;

(е) якщо $M_1 \subseteq M_2$ і M_2 — замкнена множина, то $[M_1] \subseteq M_2$.

2. З'ясувати, яке зі співвідношень $\supset, \supseteq, \not\subset, \subset, \subseteq, =$ виконується для множин K_1 і K_2 (співвідношення $\not\subset$ означає, що жодне зі співвідношень $\supset, \supseteq, \subset, \subseteq, =$ не виконується).

(а) $K_1 = [M_1 \cap M_2]$, $K_2 = [M_1] \cap [M_2]$;

(б) $K_1 = [M_1 \setminus M_2]$, $K_2 = [M_1] \setminus [M_2]$;

(в) $K_1 = [M_1 \cup (M_2 \cap M_3)]$, $K_2 = [M_1 \cup M_2] \cap [M_1 \cup M_3]$;

(г) $K_1 = [M_1 \cap (M_2 \cup M_3)]$, $K_2 = [M_1 \cap M_2] \cup [M_1 \cap M_3]$;

(д) $K_1 = [M_1 \setminus (M_1 \cap M_2)]$, $K_2 = [M_1] \setminus [M_1 \cap M_2]$.

3. Визначити систему твірних алгебри $A = \langle N_k', \{+\}\rangle$, де $N_k' = \{n \mid n \in N, n > k\}$. Довести, що ця алгебра скінченно-породжувана.

4. Довести, що алгебра $A = \langle R^n, \{+, \times\}\rangle$, де R^n — n-вимірний векторний простір, є скінченно-породжуваною ($\times$ — це операція покоординатного множення вектора з R^n на дійсне число). Вказати один із базисів алгебри A.

5. Нехай A — скінченно-породжувана алгебра. Довести, що в будь-якій системі твірних T алгебри A можна виділити скінченну підмножину $T' \subseteq T$, яка є системою твірних алгебри A.

6. Довести, що будь-яка скінченно-породжувана алгебра A має скінченний базис.

7. Довести, що будь-яка алгебра A, яка не має базису, нескінченно-породжувана.

2.6. Основні класичні алгебри

У процесі розвитку математики поступово було визначено найбільш важливі й поширені різновиди алгебр. Основою для класифікації є тип алгебри і властивості операцій її сигнатури.

Розглянемо спочатку алгебри типу (2), тобто такі, сигнатура яких містить єдину бінарну операцію.

Алгебра з єдиною бінарною операцією називається *мультиплікативною*, якщо цю операцію називають множенням, й *адитивною*, якщо операцію називають додаванням. Відповідно до цього в алгебрах використовують мультиплікативну чи адитивну термінологію. Іноді, щоб підкреслити неістотність виду операції й термінології для подальших означень, вибирають нейтральну назву й кажуть про *композицію*. Оскільки перехід від одного виду алгебри та відповідної термінології до іншого виду нескладний, зупинімося на вивченні мультиплікативних алгебр.

Півгрупа. Алгебра A з єдиною всюди означеною бінарною операцією називається ***групоїдом***.

Групоїд $A = \langle M, \{\times\}\rangle$ називається ***півгрупою***, якщо операція $\times$ асоціативна, тобто для будь-яких елементів a, b, $c \in M$ виконується рівність $(a \times b) \times c = a \times (b \times c)$.

Півгрупа $A = \langle M, \{\times\}\rangle$ називається ***комутативною***, якщо опера-ція $\times$ комутативна, тобто для довільних a, $b \in M$ виконується рівність $a \times b = b \times a$.

Півгрупа $A = \langle M, \{\times\}\rangle$ називається ***півгрупою з одиницею***, або ***моноїдом***, якщо носій M містить такий елемент e, що для всіх елементів $a \in M$ виконується рівність $a \times e = e \times a = a$. У загальному випадку елемент e називають ***нейтральним***, у мультиплікативних півгрупах — ***одиницею***, в адитивних — ***нулем***. Відзначаючи особливу роль елемента e, його можна вважати фіксованим (виділеним) елементом і відносити до сигнатури. Тоді тип півгрупи з одиницею $A = \langle M, \{\times, e\}\rangle$ — (2, 0).

Неважко довести, що коли в півгрупі є одиниця, то вона єдина. Справді, якщо припустити, що існують дві одиниці e_1 і e_2, то $e_1 \times e_2 = e_1$ і $e_1 \times e_2 = e_2$, отже $e_1 = e_2$.

Півгрупа $A = \langle M, \{\times\}\rangle$ називається ***півгрупою з правим (лівим) скороченням***, якщо для будь-яких елементів a, b, $c \in M$ з умови $a \times c = b \times c$ (відповідно, $c \times a = c \times b$) випливає рівність $a = b$.

Півгрупа A називається ***півгрупою з двостороннім скороченням*** або просто півгрупою зі скороченням, якщо вона є півгрупою з правим і лівим скороченням одночасно.

Приклад 2.7. 1. Алгебра $\langle N, \{\times\}\rangle$ — мультиплікативна комутативна півгрупа з одиницею 1. Аналогічно, алгебра $\langle N \cup \{0\}, \{+\}\rangle$ — адитивна комутативна півгрупа з нулем 0. Обидві ці півгрупи є півгрупами зі скороченням.

2. Алгебра $\langle A^*, \{^\circ\}\rangle$ — некомутативна півгрупа з одиницею. Одиницею є порожнє слово e.

3. Алгебра $\langle \beta(M), \{\cup\}\rangle$ — комутативна півгрупа з одиницею $\varnothing$.

4. Важливе місце в математиці займає некомутативна півгрупа з одиницею $A = \langle P_M, \{^\circ\}\rangle$, носієм P_M якої є множина всіх перетворень (тобто відображень у себе) множини M, а єдиною операцією — бінарна операція композиції. Неважко переконатися, що операція композиції асоціативна, а одиницею півгрупи A є тотожне перетворення $\varepsilon = i_M$.

Взагалі, для будь-якої множини перетворень $P_{M'}$ ($P_{M'} \subseteq P_M$), замкненої відносно операції композиції, алгебра $A' = \langle P_{M'}, \{^\circ\}\rangle$ буде некомутативною півгрупою. Усі такі півгрупи називають *півгрупами перетворень*, а півгрупу $A = \langle P_M, \{^\circ\}\rangle$ — *симетричною півгрупою перетворень*.

Теорема 2.3. Будь-яка півгрупа з одиницею ізоморфна деякій півгрупі перетворень.

Доведення. Нехай множина M — носій даної (мультиплікативної) півгрупи з одиницею A. Кожному елементу $a \in M$ поставимо у відповідність таке перетворення $f_a \in P_M$, що $f_a(x) = x \times a$ для всіх $x \in M$. Позначимо $P_{M'} = \{f_a \mid a \in M\}$. Замкненість $P_{M'}$ відносно операції композиції випливає з того, що $(f_a \circ f_b)(x) = f_b(f_a(x)) = f_a(x) \times b = x \times a \times b = f_{a \times b}(x)$, тобто $f_a \circ f_b = f_{a \times b} \in P_{M'}$. Отже, $\langle P_{M'}, \{^\circ\}\rangle$ — півгрупа перетворень.

Позначимо через γ означену відповідність між M і $P_{M'}$. Кожне перетворення $f_a \in P_{M'}$ за означенням множини $P_{M'}$ є образом певного елемента $a \in M$ при відображенні γ. Разом із тим, для довільних елементів a, $b \in M$, $a \neq b$, виконуються співвідношення $f_a(e) = a \neq b = f_b(e)$, тобто $f_a \neq f_b$. Отже, відповідність γ — всюди визначена, функціональна, сюр'єктивна й ін'єктивна, тобто є взаємно однозначним відображенням.

Зі співвідношення $\gamma(a \times b) = f_{a \times b} = f_a \circ f_b = \gamma(a) \circ \gamma(b)$ випливає справедливість умови (2.1). Отже, γ — ізоморфізм півгрупи $\langle M, \{\times\}\rangle$ на півгрупу $\langle P_{M'}, \{^\circ\}\rangle$.

Група. Півгрупа з одиницею $A = \langle M, \{\times\}\rangle$ називається ***групою***, якщо для будь-яких a, $b \in M$ у множині M завжди існують такі елементи x і y, що $x \times a = b$ і $a \times y = b$.

Існує інше *рівносильне означення* поняття групи: півгрупа з одиницею $A = \langle M, \{\times\}\rangle$ називається *групою*, якщо для будь-якого $a \in M$ існує такий елемент $b \in M$, що $a \times b = b \times a = e$. Елемент b називають ***оберненим*** (або ***симетричним***) до елемента a та позначають a^{-1}.

Можна довести, що для кожного елемента $a \in M$ групи A існує тільки один обернений елемент, а також що для елементів групи виконуються умови правого і лівого скорочення. Крім того, для довільних елементів a, $b \in M$ групи A виконуються співвідношення $(a^{-1})^{-1} = a$ й $(a \times b)^{-1} = b^{-1} \times a^{-1}$. Доведення всіх наведених тверджень, а також доведення рівносильності обох означень групи пропонуємо провести самостійно.

Група $A = \langle M, \{\times\} \rangle$ називається ***комутативною***, чи ***абелевою*** групою, якщо операція $\times$ комутативна.

Приклад 2.8. 1. Алгебра $\langle Z, \{+\} \rangle$ є комутативною групою. Її нейтральним елементом є число 0, а оберненим елементом a^{-1} для довільного елемента $a \in Z$ — число $-a$.

2. Алгебра $\langle Q\backslash\{0\}, \{\times\} \rangle$ є комутативною групою. Оберненим до елемента a є число $1/a$, а одиницею групи — число 1.

3. Розглянемо множину $N_{(p)} = \{0, 1, 2, \ldots, p-1\}$ і операцію $\oplus$ додавання за модулем p (див. приклад 2.2.2). Тоді $\langle N_{(p)}, \{\oplus\} \rangle$ — комутативна група, нейтральним елементом якої є число 0, а оберненим елементом для числа $k \in N_{(p)}$ — є число $p - k$ (оберненим елементом для числа 0 є 0).

4. Алгебра $\Pi = \langle P_M, \{\circ, ^{-1}, \varepsilon\} \rangle$ із прикладу 2.2.2 є групою, яку називають *симетричною групою підстановок*. Нехай P' — деяка замкнена відносно операції композиції сукупність підстановок множини M ($P' \subseteq P_M$), для кожного елемента φ якої існує обернений елемент $\varphi^{-1} \in P'$. Тоді P' разом з операцією композиції утворює групу, яку називають *групою підстановок*. Слід зауважити, що умова скінченності множини M неістотна й, отже, можна розглядати групи підстановок для довільної множини M.

Аналогічно теоремі 2.3 можна довести таку теорему.

Теорема 2.4 (*Келі*). Довільна група $A = \langle M, \{\times\} \rangle$ ізоморфна деякій групі підстановок.

Доведення. Як і раніше, кожному елементу $a \in M$ поставимо у відповідність таке перетворення φ_a носія M у себе, що $\varphi_a(x) = x \times a$ для всіх $x \in M$. Доведемо, що перетворення φ_a — підстановка. Усюди визначеність і функціональність φ_a випливає відповідно з усюди визначеності та функціональності операції $\times$. Покажемо сюр'єктивність відображення φ_a. Справді, кожен елемент $c \in M$ є образом деякого елемента $b \in M$ при відображенні φ_a, а саме $b = c \times a^{-1}$, бо $\varphi_a(c \times a^{-1}) = (c \times a^{-1}) \times a = c$. Відображення φ_a — ін'єктивне, оскільки, якщо $\varphi_a(x) = \varphi_a(x')$, то $x \times a = x' \times a$ й $(x \times a) \times a^{-1} = (x' \times a) \times a^{-1}$, тобто $x = x'$. Отже, φ_a — бієкція.

Покладемо $P' = \{\varphi_a \mid a \in M\}$ і доведемо, що $A' = \langle P', \{\circ\} \rangle$ — група підстановок. Замкненість множини підстановок P' відносно операції композиції випливає з того, що $\varphi_a \circ \varphi_b = \varphi_{a \times b} \in P'$, а існування для кожної підстановки $\varphi_a \in P'$ оберненої — з того, що $\varphi_a^{-1} = \varphi_{a^{-1}} \in P'$.

Позначимо через γ означену вище відповідність між M і P'. Всюди визначеність, функціональність і сюр'єктивність γ випливають безпосередньо з означення. Крім того, для довільних елементів a, $b \in M$, $a \neq b$, маємо $\varphi_a(e) = a \neq b = \varphi_b(e)$, тобто $\varphi_a \neq \varphi_b$. Таким чином, відповідність γ є також ін'єктивною, а отже, взаємно однозначним відображенням. Нарешті, маємо $\gamma(a \times b) = \varphi_{a \times b} = \varphi_a \circ \varphi_b = \gamma(a) \circ \gamma(b)$. Таким чином, γ — ізоморфізм алгебр A й A'.

Наведені нижче алгебри мають тип (2, 2), тобто є алгебрами з двома бінарними операціями. Одну з цих операцій прийнято називати *додаванням* і позначати +, а другу — *множенням* (знак $\times$).

Кільце і поле. Алгебру $\langle M, \{+, \times\} \rangle$ називають ***півкільцем***, якщо $\langle M, \{+\} \rangle$ — комутативна півгрупа зі скороченням, $\langle M, \{\times\} \rangle$ — півгрупа, а операції + і $\times$ пов'язані законом дистрибутивності, тобто $(a + b) \times c = a \times c + b \times c$ для довільних a, b, $c \in M$.

Якщо півгрупа $\langle M, \{\times\} \rangle$ комутативна, то півкільце називають *комутативним*.

Півкільце $K = \langle M, \{+, \times\} \rangle$ називають ***кільцем***, якщо його адитивна півгрупа є групою.

Кожне кільце містить нейтральний елемент 0 відносно операції додавання. Якщо кільце містить нейтральний елемент 1 відносно операції множення, то його називають ***кільцем з одиницею***.

Кільце з одиницею $\langle M, \{+, \times\} \rangle$ називають ***тілом***, якщо множина M містить принаймні два елементи і $\langle M\backslash\{0\}, \{\times\} \rangle$ — група.

Тіло $\langle M, \{+, \times\} \rangle$ називають ***полем***, якщо група $\langle M\backslash\{0\}, \{\times\} \rangle$ комутативна.

Решітка. Нарешті, розглядувану вище алгебричну систему — решітку — можна означити, не використовуючи відношення часткового порядку, як алгебру типу (2, 2). При означенні решітки адитивну операцію зазвичай позначають $\vee$ і називають *об'єднанням*, а мультиплікативну операцію — $\wedge$ і називають *перетином*.

Алгебра $L = \langle M, \{\vee, \wedge\} \rangle$ називається ***решіткою***, якщо операції $\vee$ і $\wedge$ всюди визначені, асоціативні, комутативні та для будь-яких елементів a, $b \in M$ виконуються співвідношення:

$$a \vee a = a,\ a \wedge a = a \qquad (\textit{ідемпотентність}),$$
$$a \vee (a \wedge b) = a,\ a \wedge (a \vee b) = a \quad (\textit{правила поглинання}). \qquad (2.3)$$

Можна довести, що наведене означення й означення решітки в прикладі 2.1.2 рівносильні між собою. Потрібне відношення частково-го порядку на множині M можна означити так: для $a, b \in M$ вважаємо $a \le b$ тоді й тільки тоді, коли $a \wedge b = a$ й $a \vee b = b$. Тоді $\inf\{a, b\} = a \wedge b$ і $\sup\{a, b\} = a \vee b$.

Якщо для довільних елементів $a, b, c \in M$ решітки $L = \langle M, \{\vee, \wedge\}\rangle$ виконуються також умови $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$, $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, то решітка називається ***дистрибутивною***.

Решітка $L = \langle M, \{\vee, \wedge\}\rangle$ називається ***решіткою з нулем і одиницею***, якщо в множині M існують елементи, традиційно позначувані як **0** і **1**, такі, що для всіх $a \in M$ виконуються співвідношення

$$\mathbf{0} \wedge a = \mathbf{0},\ \mathbf{0} \vee a = a,\ \ \mathbf{1} \wedge a = a,\ \mathbf{1} \vee a = \mathbf{1}. \qquad (2.4)$$

Приклад 2.9. 1. Алгебри $\langle Z, \{+, \times\}\rangle$, $\langle Z_{(k)}, \{+, \times\}\rangle$ і $\langle N_{(p)}, \{\oplus, \otimes\}\rangle$ — комутативні кільця з одиницею (див. приклади 2.2 та 2.3).

2. Алгебри $\langle Q, \{+, \times\}\rangle$, $\langle R, \{+, \times\}\rangle$ та $\langle C, \{+, \times\}\rangle$ є полями.

3. Алгебра $P = \langle B, \{+, \times\}\rangle$, де $B = \{0, 1\}$ є полем, якщо операції + і $\times$ означити так: $0 + 0 = 1 + 1 = 0$, $0 + 1 = 1 + 0 = 1$ і $0 \times 0 = 0 \times 1 = 1 \times 0 = 0$, $1 \times 1 = 1$. Покладемо також $-0 = 0$, $-1 = 1$ й $1^{-1} = 1$. Поле P — скінченне, причому його носій містить найменшу можливу кількість елементів.

4. Алгебра множин $L = \langle \beta(M), \{\cup, \cap, {}^{-}\}\rangle$ — дистрибутивна решітка з нулем і одиницею. Нулем решітки R є порожня множина $\varnothing$, а одиницею — множина M.

2.5. Задачі і вправи

1. Довести, що алгебра $A = \langle \beta(M), \{\cap\}\rangle$ — комутативна півгрупа з одиницею. Чи є A півгрупою зі скороченням?

2. Довести, що алгебра $A = \langle V, \{\circ\}\rangle$, де V — множина всіх бінарних відношень на множині M і $\circ$ — операція композиції відношень, є некомутативною півгрупою з одиницею.

3. Довести, що алгебра $A = \langle Z, \{*\}\rangle$, у якій операцію * означено рівністю $n * m = n + m + nm$, — комутативний моноїд.

4. Довести, що алгебра $A = \langle N, \{*\}\rangle$, у якій операцію * означено співвідношенням $n * m = \max(n, m)$, є комутативним моноїдом. Чи буде A півгрупою зі скороченням?

5. Довести, що алгебра $A = \langle \{a, b\}, \{*\}\rangle$, у якій операцію $*$ означено співвідношеннями $a * a = a$, $a * b = b * a = b * b = b$, є півгрупою. Чи є ця півгрупа

(а) комутативною; (б) моноїдом; (в) півгрупою зі скороченням?

6. Чи є півгрупою алгебра $A = \langle \{a, b\}, \{*\}\rangle$, у якій операцію $*$ означено співвідношеннями $a * a = b$, $a * b = b * a = b * b = a$?

7. Довести рівносильність обох наведених вище означень групи.

8. Довести, що для будь-якого елемента групи існує тільки один обернений елемент.

9. Довести, що для елементів довільної групи виконуються умови правого та лівого скорочення.

10. Довести, що для всіх елементів $a, b \in M$ групи $G = \langle M, \{\times\}\rangle$ виконується рівність $(a \times b)^{-1} = b^{-1} \times a^{-1}$.

11. Довести, що в мультиплікативній групі не можуть існувати рівно два такі неодиничні елементи a та b, що $a \times a = e$ й $b \times b = e$.

12. Навести приклад скінченної півгрупи з лівим скороченням, яка не є групою.

13. Побудувати групу, носієм якої є трьохелементна множина $M = \{e, a, b\}$.

14. На множині $M = (R \backslash \{0\}) \times R$ означимо операцію множення $*$: $(a, b) * (c, d) = (ac, ad + b)$. Довести, що $A = \langle M, \{*\}\rangle$ — некомутативна група.

15. Нехай для будь-якого елемента $a \in M$ групи $G = \langle M, \{\times\}\rangle$ виконується $a \times a = e$. Довести, що G є комутативною групою.

16. Довести, що алгебра $A = \langle M, \Omega\rangle$ є комутативною групою:

(а) $M = \{ -1, 1 \}$, $\Omega = \{\times\}$;

(б) $M = \{ -1, 1, i, -i \}$, $\Omega = \{\times\}$;

(в) M — множина матриць виду $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $a, b \in R$, $\Omega = \{+\}$;

(г) M — множина матриць виду $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $a,\ b \in R$ і $a^2 + b^2 \ne 0$,

а Ω містить операцію множення матриць.

17. Довести, що алгебра $A = \langle N \times N, \{+, \times\}\rangle$, в якій операції + і $\times$ означено рівностями $(n, m) + (k, l) = (n + k, m + l)$, $(n, m) \times (k, l) = (nk + ml, nl + mk)$, є комутативним півкільцем.

18. Довести, що алгебра $A = \langle N \times N, \{+, \times\}\rangle$, в якій операції + і $\times$ означено рівностями $(n, m) + (k, l) = (n + k, m + l)$, $(n, m) \times (k, l) = (nk, ml)$, є комутативним півкільцем з одиницею.

19. Довести, що для довільних елементів $a, b \in M$ кільця $K = \langle M, \{+, \times\}\rangle$ виконується така рівність:

(а) $a \times 0 = 0 \times a = 0$; (в) $(-a) \times b = a \times (-b)$;

(б) $(-1) \times a = -a$; (г) $(-a) \times (-b) = a \times b$.

20. Нехай F — множина всіх відображень типу $R \to R$. Означимо на F операції + та × рівностями $(f + g)(x) = f(x) + g(x)$, $(f \times g)(x) = f(x) \times g(x)$. Чи є кільцем алгебра:

(а) $\langle F, \{+, \times\}\rangle$;

(б) $\langle F, \{+, \circ\}\rangle$, де $\circ$ — операція композиції?

21. Чи є кільцем алгебра $A = \langle \beta(M), \{\cup, \cap\}\rangle$?

22. Довести, що алгебра $A = \langle M, \{+, \times\}\rangle$, де $M = \{a + b\sqrt{2} \mid a, b \in Q\}$, є полем.

23. Чи є алгебра остач $A = \langle N_{(p)}, \{\oplus, \otimes\}\rangle$ полем для

(а) $p = 4$; (б) $p = 3$; (в) $p = 2$?

24. Довести, що алгебра остач $A = \langle N_{(p)}, \{\oplus, \otimes\}\rangle$ є кільцем, а коли p — просте число, то полем.

25. Довести, що алгебра $A = \langle R^2, \{+, \times\}\rangle$, операції якої означено рівностями $(a, b) + (c, d) = (a + c, b + d)$, $(a, b) \times (c, d) = (ac - bd, ad + bc)$, є полем.

26. Довести, що поле $A = \langle R^2, \{+, \times\}\rangle$ з попереднього прикладу ізоморфне полю комплексних чисел $\langle C, \{+, \times\}\rangle$.

27. Довести, що для довільних елементів $a, b \in M$ решітки $A = \langle M, \{\vee, \wedge\}\rangle$ виконується

(а) $a \vee b = a \wedge b$ тоді й тільки тоді, коли $a = b$;

(б) $a \vee b = b$ тоді й тільки тоді, коли $a \wedge b = a$.

28. Довести рівносильність обох вищенаведених означень решітки.

29. Довести, що алгебра $A = \langle D, \{\vee, \wedge\}\rangle$, де D — множина всіх дільників певного числа $n \in N$, а операції $\vee$ та $\wedge$ означено співвідношеннями $k \vee m =$ НСК(k, m) і $k \wedge m =$ НСД(k, m), є дистрибутивною решіткою з нулем і одиницею.

2.7. Булева алгебра

Окремо розглянемо один з найважливіших і найпоширеніших різновидів алгебр — так звану булеву алгебру, або алгебру Буля.

Булевою алгеброю (***булевою решіткою***) називається алгебра $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ типу (2, 2, 1, 0, 0), сигнатура якої містить дві бінарні операції додавання $\vee$ та множення $\wedge$, унарну операцію заперечення $\neg$ та два виділені елементи (нульарні операції) нуль **0** і одиницю **1**,

якщо для довільних елементів a, b, $c \in M$ мають місце такі співвідношення:

1. $(a \vee b) \vee c = a \vee (b \vee c)$, $(a \wedge b) \wedge c = a \wedge (b \wedge c)$, (*асоціативність*)
2. $a \vee b = b \vee a$, $a \wedge b = b \wedge a$, (*комутативність*),
3. $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$,
   $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, (*дистрибутивність*),
4. $a \vee \mathbf{0} = a$, $a \vee \mathbf{1} = \mathbf{1}$, (*властивості нуля й одиниці*),
5. $a \wedge \mathbf{0} = \mathbf{0}$, $a \wedge \mathbf{1} = a$,
6. $a \vee \neg a = \mathbf{1}$, $a \wedge \neg a = \mathbf{0}$. (*властивості доповнення*).

Вимоги, або умови 1–6 називають ***аксіомами булевої алгебри***.

Будь-яка булева алгебра $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ є дистрибутивною решіткою з нулем і одиницею. Асоціативність, комутативність і дистрибутивність операцій $\vee$ та $\wedge$ й властивості (2.4) елементів **0** та **1** відповідної решітки випливають безпосередньо з аксіом 1–5 булевої алгебри. Властивості ідемпотентності та поглинання решітки (2.3) можна довести за допомогою таких перетворень:

$a = a \vee \mathbf{0} = a \vee (a \wedge \neg a) = (a \vee a) \wedge (a \vee \neg a) = (a \vee a) \wedge \mathbf{1} = a \vee a$,

$a = a \wedge \mathbf{1} = a \wedge (a \vee \neg a) = (a \wedge a) \vee (a \wedge \neg a) = (a \wedge a) \vee \mathbf{0} = a \wedge a$,

$a \vee (a \wedge b) = (a \wedge \mathbf{1}) \vee (a \wedge b) = a \wedge (b \vee \mathbf{1}) = a \wedge \mathbf{1} = a$,

$a \wedge (a \vee b) = (a \vee \mathbf{0}) \wedge (a \vee b) = a \vee (\mathbf{0} \wedge b) = a \vee \mathbf{0} = a$.

Беручи до уваги зазначений зв'язок між булевими алгебрами та решітками, іноді пропонують інше рівносильне означення булевої алгебри. *Булевою алгеброю* називають дистрибутивну решітку з нулем і одиницею, у якій для кожного елемента a додатково означено унарну операцію *заперечення* $\neg a$ таку, що $a \vee \neg a = \mathbf{1}$ й $a \wedge \neg a = \mathbf{0}$.

У наведеному загальному означенні булевої алгебри для позначення операцій сигнатури використано найбільш популярні й загальновживані символи. У кожному конкретному випадку булевої алгебри ці позначення та назви операцій можуть змінюватися. Наприклад, операцію додавання $\vee$ називають також *диз'юнкцією*; крім того, для неї використовують назви *об'єднання* та *точна верхня грань* і відповідні позначення $\cup$ і sup. Аналогічно, операцію множення $\wedge$ в деяких булевих алгебрах називають *кон'юнкцією*, *перетином* або *точною нижнею гранню* та позначають відповідно &, $\cap$ або inf. Операцію заперечення $\neg$ називають іноді операцією *доповнення*, і для неї використовують позначення a' або $\bar{a}$.

Класичним прикладом булевої алгебри є означена в прикладі 2.2.4 булева алгебра множин $A = \langle \beta(M), \{\cup, \cap, {}^{-}, \varnothing, M\}\rangle$ над довільною множиною M. Інші важливі приклади булевих алгебр — *алгебра висловлювань*, *алгебра релейно-контактних схем*, *алгебра комбінаційних (вентильних, перемикальних) схем*, *алгебра булевих функцій*, *алгебра регулярних подій* і т. ін. Кожну з цих алгебр ми означимо та детальніше вивчатимемо далі.

2.6. Задачі і вправи

1. Нехай B^n — множина двійкових кортежів довжиною n, а операції $\vee$ та $\wedge$ для елементів із B^n означено як у розділі 2.3. Довести, що $A = \langle B^n, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ — булева алгебра, де $\mathbf{0} = (0, 0, \ldots, 0)$ і $\mathbf{1} = (1, 1, \ldots, 1)$.

2. Нехай $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ — булева алгебра. Ненульовий елемент $a \in M$ називається *атомом*, якщо з $b \wedge a = b$ випливає $b = \mathbf{0}$ чи $b = a$. Довести, що для скінченної булевої алгебри істинне таке твердження:

(а) якщо $x \neq \mathbf{0}$, то існує атом a, для якого $a \wedge x = a$;

(б) якщо a — атом та $x \in M$, то виконується одне й тільки одне зі співвідношень: $a \wedge x = a$ чи $a \wedge x = \mathbf{0}$;

(в) якщо a — атом та $x \in M$, то виконується одне й тільки одне зі співвідношень: $a \wedge x = a$ чи $a \wedge (\neg x) = a$.

3. Нехай $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ — булева алгебра. Для $x \in M$ через $\mathrm{At}(x)$ позначимо множину всіх атомів a алгебри A, для яких $a \wedge x = a$. Довести, що для скінченної булевої алгебри A справедливе таке твердження:

(а) для будь-яких $x, y \in M$ $\mathrm{At}(x \wedge y) = \mathrm{At}(x) \cap \mathrm{At}(y)$;

(б) для будь-яких $x, y \in M$ $\mathrm{At}(x \vee y) = \mathrm{At}(x) \cup \mathrm{At}(y)$;

(в) для будь-якого $x \in M$ $\mathrm{At}(\neg x) = \mathrm{At}(\mathbf{1}) \setminus \mathrm{At}(x)$;

(г) $\mathrm{At}(x) = \mathrm{At}(y)$ тоді й тільки тоді, коли $x = y$;

(д) якщо a — атом, то $\mathrm{At}(a) = \{a\}$;

(е) якщо $a_1, a_2, \ldots, a_k$ — атоми, то $\mathrm{At}(a_1 \vee a_2 \vee \ldots \vee a_k) = \{a_1, a_2, \ldots, a_k\}$.

4. Довести, що довільна нетривіальна скінченна булева алгебра $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ ($|M|>1$) ізоморфна булевій алгебрі множин $A' = \langle \beta(L), \{\cup, \cap, ^{-}, \varnothing, L\}\rangle$ для деякої скінченної множини L (*теорема Стоуна*).

5. Довести, що довільні скінченні булеві алгебри з однаковою кількістю елементів ізоморфні між собою.

6. Довести, що кількість елементів у будь-якій скінченній булевій алгебрі є степенем 2.

7. Побудувати булеві алгебри з двома, чотирма та вісьмома елементами.

8. Довести, що коли в булевій алгебрі $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ для деякого $a \in M$ існує такий елемент b, що $a \vee b = \mathbf{1}$ та $a \wedge b = \mathbf{0}$, то $b = \neg a$.

9. Довести, що в булевій алгебрі $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ для довільних елементів $a, b \in M$ виконується рівність:

(а) $\neg(\neg a) = a$;

(б) $\neg\mathbf{0} = \mathbf{1}$, $\neg\mathbf{1} = \mathbf{0}$;

(в) $\neg(a \vee b) = (\neg a) \wedge (\neg b)$, $\neg(a \wedge b) = (\neg a) \vee (\neg b)$.

10. Довести, що в булевій алгебрі такі чотири співвідношення рівносильні:

(а) $a \wedge b = a$; (в) $a \wedge (\neg b) = \mathbf{0}$;

(б) $a \vee b = b$; (г) $\neg a \vee b = \mathbf{1}$.

11. Довести, що в булевій алгебрі $A = \langle M, \{\vee, \wedge, \neg, \mathbf{0}, \mathbf{1}\}\rangle$ для елементів $a, b \in M$ рівність $((a \wedge (\neg b)) \vee ((\neg a) \wedge b)) = \mathbf{0}$ виконується тоді й тільки тоді, коли $a = b$.

РОЗДІЛ 3

МАТЕМАТИЧНА ЛОГІКА

Математична логіка як самостійний розділ сучасної математики сформувалася відносно нещодавно — на рубежі XIX та XX століть. Виникнення та швидкий розвиток математичної логіки були пов'язані з так званою кризою основ математики, одним із проявів якої є відомі парадокси, або антиномії, канторівської теорії множин (див. розділ 1.14).

Основним предметом у дослідженнях, присвячених "ліквідуванню" кризи та "рятуванню" математики, стали принципи (правила) побудови математичних тверджень і теорій, зокрема пошук відповіді на такі питання: "Як має бути побудована теорія, щоб у ній не виникало суперечностей (антиномій)?", "Які властивості повинні мати методи доведення, щоб їх можна було вважати строгими?" тощо.

У математиці з античних часів існував зразок систематичної і строгої побудови теорії — *геометрія Евкліда*, у якій усі початкові положення сформульовано явно, у вигляді аксіом, а всі твердження, істинні в цій теорії, — теореми — можна вивести з цих аксіом за допомогою послідовностей логічних міркувань, що називаються доведеннями.

Однак при побудові більшості наступних математичних теорій математики зазвичай не вважали за потрібне явно виділяти всі вихідні принципи та чітко формулювати методи конструювання доведень; критерії строгості доведень і очевидності тверджень у математиці в різні часи були різними. Відтак це призводило час від часу до виникнення криз і потреби перегляду основ тієї чи іншої теорії.

3.1. Принципи побудови формальних теорій

Наприкінці XIX століття у зв'язку з виникненням кризи в канторівській теорії множин постала потреба переглянути загальні принципи організації математичних теорій. Це зумовило створення нової галузі математики — *засад математики.*

Однією з фундаментальних ідей, на які спираються дослідження із засад математики, є ідея *формалізації теорії*, тобто послідовного застосування аксіоматичного методу побудови теорії. При цьому не можна використовувати будь-які припущення про об'єкти теорії, окрім тих, що виражені явно у вигляді аксіом. Аксіоми розглядають як формальні послідовності символів (вирази, формули або слова), а методи доведення — як методи одержання одних виразів з інших за допомогою операцій над символами.

Такий формальний алгебричний підхід гарантує чіткість і однозначність вихідних (початкових) тверджень та коректність і однозначність виводу. Однак може скластися враження, що осмисленість (зміст, інтерпретація чи семантика) понять і тверджень у формалізованій теорії не відіграють жодної ролі. Зовні це так і є; однак насправді й аксіоми, і правила виводу прагнуть означати так, щоб побудована за їх допомогою формальна теорія була б змістовною.

У найзагальнішому вигляді ***формальну теорію*** T (***числення***) будують таким чином:

1. Означають набір основних символів — ***алфавіт*** теорії.
2. Конструктивно (як правило, індуктивно) означають ***множину формул***, або правильно побудованих виразів, яка утворює мову теорії.
3. Виокремлюють підмножину формул, які називають ***аксіомами*** теорії.
4. Задають ***правила виводу (виведення)*** теорії.

Правило виводу $R(F_1, F_2, \ldots, F_m, G)$ — це відношення (або операція) на множині формул.

Якщо формули $F_1, F_2, \ldots, F_m, G$ перебувають у відношенні R, то формула G називається ***безпосередньо вивідною*** з формул $F_1, F_2, \ldots, F_m$ за правилом R.

Часто правило виводу $R(F_1, F_2, \ldots, F_m, G)$ записують у вигляді

$$\frac{F_1, F_2, \ldots, F_m}{G}.$$

Формули $F_1, F_2, \ldots, F_m$ називають ***припущеннями***, ***посилками*** або ***гіпотезами*** правила R, а формулу G — ***висновком***, ***наслідком*** або ***вислідом***.

Виведенням (***виводом***) формули B з формул $A_1, A_2, \ldots, A_n$ називають таку послідовність формул $F_1, F_2, \ldots, F_m$, що $F_m = B$, а будь-яка формула F_i, $i = 1, 2, \ldots, m$, є:

1) або аксіомою;

2) або однією з початкових формул $A_1, A_2, \ldots, A_n$;

3) або формулою безпосередньо вивідною з формул $F_1, F_2, \ldots, F_{i-1}$ (або будь-якої їх підмножини) за одним із правил виводу.

Якщо існує виведення формули B з формул $A_1, A_2, \ldots, A_n$, то кажуть, що B є ***вивідною*** з $A_1, A_2, \ldots, A_n$ і позначають цей факт так: $A_1, A_2, \ldots, A_n \vdash B$. Формули $A_1, A_2, \ldots, A_n$ називають ***посилками***, або ***гіпотезами***, виведення. Перехід у виведенні від формули F_{i-1} до F_i називають i-м кроком виведення.

Доведенням формули B в теорії T називають виведення B з порожньої множини формул, тобто виведення, у якому як початкові формули використовують тільки аксіоми теорії.

Формула B, для якої існує доведення, називається формулою, ***довідною*** (***вивідною***) у теорії T, або ***теоремою*** теорії T; факт довідності формули B позначають $\vdash B$.

При вивченні формальних теорій слід розрізняти два типи тверджень:

1) твердження самої теорії, або її теореми;

2) твердження про теорію (про властивості її теорем, доведень тощо).

Перші є елементами (словами, виразами, формулами) внутрішньої мови теорії, а другі — зовнішніми; останні формулюють у термінах мови, зовнішньої щодо теорії і званої ***метамовою*** теорії; самі ці твердження називають ***метатеоремами***. Наприклад, якщо побудовано доведення формули B, то вираз B — теорема, а твердження "$\vdash B$" — метатеорема.

3.2. Алгебра висловлень

Носієм алгебри висловлень є множина так званих простих висловлень.

Просте (***елементарне***) ***висловлення*** (*висловлювання*) — це просте твердження, тобто розповідне речення, щодо змісту якого доречно ставити питання про його правильність або неправильність.

Прості висловлення, у яких виражено правильну думку, називатимемо ***істинними***, а ті, що виражають неправильну думку, — ***хибними***.

Поняття простого висловлення, поняття істинності та хибності належать до первинних неозначуваних понять математики, таких як "число", "пряма", "точка", "площина", "множина" тощо.

Наведемо кілька прикладів елементарних висловлень:

1) Київ — столиця України;

2) число 7 — просте;

3) число 10 більше від числа 3;

4) усі натуральні числа — прості;

5) множина всіх простих чисел — скінченна.

Перші три висловлення істинні, а два останні — хибні.

Водночас речення "Хай живе математична логіка!" або "Уважно прочитайте весь цей розділ" — не є висловленнями.

Розглядаючи висловлення, виходитимемо з двох основних припущень:

1) кожне висловлення є або істинним, або хибним (***закон виключення третього***);

2) жодне висловлення не є одночасно істинним і хибним (***закон виключення суперечності***).

Приймаючи ці припущення, ми стаємо на позиції класичної (традиційної) двозначної логіки. У XX столітті виникли і продовжують досліджуватись так звані некласичні логіки: багатозначна, інтуїціоністська (конструктивна), модальна, нечітка тощо. У подальшому ми додержуватимемося принципів класичної логіки, в рамках якої проводитимуться всі математичні міркування.

Позначатимемо елементарні висловлення малими латинськими літерами: $a, b, c, \ldots$ (можливо, з індексами), а значення висловлень "Істинно" і "Хибно" — відповідно символами 1 і 0 або I і X.

Крім того, розглядатимемо так звані ***змінні висловлення***, які позначатимемо латинськими літерами $x, y, z, \ldots$ (можливо, з індексами) і називатимемо також ***пропозиційними змінними***. Після підстановки замість пропозиційної змінної певного елементарного висловлення ця змінна набуде відповідного значення: 0 або 1.

Сигнатура алгебри висловлень традиційно складається з таких операцій: ***заперечення***, ***кон'юнкція***, ***диз'юнкція*** й ***імплікація***.

Різні назви і позначення, що використовують для зазначених операцій, наведено в табл. 3.1.

Таблиця 3.1

Назва	Позначення
Кон'юнкція (логічне множення, логічне "І")	$\wedge$ &
Диз'юнкція (логічне додавання, логічне "АБО")	$\vee$
Заперечення (логічне "НІ")	$\neg$
Імплікація (логічне слідування)	$\rightarrow$ $\supset$

Використовуватимемо перші з наведених назв і позначень. Нижче подано табл. 3.2, що містить означення цих операцій.

Таблиця 3.2

x y	$x \wedge y$	$x \vee y$	$\neg x$	$x \rightarrow y$
0 0	0	0	1	1
0 1	0	1	1	1
1 0	0	1	0	0
1 1	1	1	0	1

Застосовуючи до елементарних висловлень і пропозиційних змінних означені операції, діставатимемо ***складені висловлення***, яким відповідатимуть так звані ***формули***, або вирази алгебри висловлень. Для запису цих формул, дослідження їх властивостей і співвідношень між формулами та висловленнями використовують *формальні мови*, тобто певні множини слів у деякому алфавіті.

Алфавіт найпоширенішої формальної мови алгебри висловлень містить такі три групи символів:

1) символи елементарних висловлень і пропозиційних змінних: $a, b, c, \ldots$ й $x, y, z, \ldots$ (можливо, з індексами);

2) символи операцій: $\wedge, \vee, \neg, \rightarrow$;

3) допоміжні символи — круглі дужки: (і).

Із пропозиційних змінних і елементарних висловлень за допомогою операцій і дужок будують пропозиційні формули, або просто ***формули*** алгебри висловлень за такими індуктивними правилами:

1) усі пропозиційні змінні й елементарні висловлення є формулами;

2) якщо A і B — формули, то вирази $(A \wedge B)$, $(A \vee B)$, $(\neg A)$, $(A \rightarrow B)$ — також формули;

3) інших формул, ніж побудовані за правилами 1) і 2), немає.

Традиційно формули алгебри висловлень позначають великими готичними літерами, але для зручності позначатимемо їх великими латинськими літерами.

Кожна формула A зображує форму, або схему, складеного висловлення: вона перетворюється у висловлення, якщо замість її пропозиційних змінних підставити будь-які висловлення. Оскільки кожне з підставлених висловлень має значення 0 або 1, то, послідовно обчислюючи значення всіх підформул формули A, одержимо значення формули A на цьому наборі висловлень, яке дорівнює 0 або 1. Підставляючи у формулу A замість її пропозиційних змінних інший набір висловлень, аналогічно обчислимо нове значення формули A і т. д. Оскільки кожне з висловлень набору повністю характеризується своїм значенням (істинно чи хибно, тобто 1 чи 0), то замість пропозиційних змінних у формулу можна підставляти не самі висловлення, а їх значення — 1 або 0.

Нехай $p_1, p_2, \ldots, p_n$ — усі пропозиційні змінні, що входять у формулу A; позначатимемо це $A(p_1, p_2, \ldots, p_n)$. Формулі $A(p_1, p_2, \ldots, p_n)$ поставимо у відповідність функцію $f(p_1, p_2, \ldots, p_n)$, що означена на множині B^n ($B = \{0, 1\}$) і набуває значення у множині B, тобто функцію типу $B^n \rightarrow B$. Значення функції f на наборі значень $a_1, a_2, \ldots, a_n$ її змінних $p_1, p_2, \ldots, p_n$ дорівнює значенню формули $A(p_1, p_2, \ldots, p_n)$ при підстановці в неї замість пропозиційних змінних $p_1, p_2, \ldots, p_n$ відповідно значень $a_1, a_2, \ldots, a_n$.

Зауважимо, що кількість елементів в області визначення функції f дорівнює 2^n, тобто $|\mathrm{Pr}_1 f| = 2^n$.

Функцію f називають ***функцією істинності*** для формули A чи відповідного складеного висловлення. Для функції істинності можна побудувати так звану ***таблицю істинності*** (табл. 3.3).

Серед формул алгебри висловлень особливе місце займають ті формули $A(p_1, p_2, \ldots, p_n)$, які на всіх наборах $(a_1, a_2, \ldots, a_n)$ значень своїх змінних набувають значення 1.

Формулу алгебри висловлень $A(p_1, p_2, \ldots, p_n)$ називають ***тавтологією***, коли їй відповідає функція істинності, що тотожно дорівнює 1.

Тавтології ще називають *тотожно істинними* формулами, або *законами* алгебри висловлень. Аналогом тавтології у природній мові є поняття істинного твердження.

Таблиця 3.3

$p_1 p_2 \dots p_{n-1} p_n$	$f(p_1, p_2, \dots, p_{n-1}, p_n)$
0 0 … 0 0	$f(0, 0, \dots, 0, 0)$
0 0 … 0 1	$f(0, 0, \dots, 0, 1)$
0 0 … 1 0	$f(0, 0, \dots, 1, 0)$
0 0 … 1 1	$f(0, 0, \dots, 1, 1)$
………………	………………
1 1 … 1 0	$f(1, 1, \dots, 1, 0)$
1 1 … 1 1	$f(1, 1, \dots, 1, 1)$

Наведемо приклади деяких важливих тавтологій:

$(p \vee (\neg p))$ (*закон виключення третього*),

$(\neg(p \wedge (\neg p)))$ (*закон виключення суперечності*),

$(p \rightarrow p)$ (*закон тотожності*).

Довести, що ці формули є тавтологіями, можна за допомогою відповідних таблиць істинності. Те, що формула A алгебри висловлень є тавтологією, позначають так: $\models A$. Символ $\models$, як і A, належать метамові.

Формула алгебри висловлень $A(p_1, p_2, \dots, p_n)$, яка набуває значення 0 на всіх наборах $(a_1, a_2, \dots, a_n)$ значень своїх пропозиційних змінних, називається ***суперечністю***, або *тотожно хибною* формулою.

Формулу, яка не є ні тавтологією, ні суперечністю, називають ***нейтральною***.

Множину всіх формул алгебри висловлень можна розбити на тавтології, суперечності та нейтральні формули.

Формула, яка не є суперечністю, називається ***виконуваною***.

Наведемо ряд тверджень, які неважко об´рунтувати.

1. Заперечення тавтології є суперечністю, і навпаки.

2. Кожна тавтологія є виконуваною формулою (обернене твердження неправильне).

3. Кожна нейтральна формула виконувана, але не навпаки.

4. Заперечення виконуваної формули може бути як виконуваною, так і невиконуваною формулою.

Дві формули A і B алгебри висловлень називають ***рівносильними*** (або *логічно еквівалентними*), якщо їм відповідає та сама функція істинності. Рівносильність формул A і B позначають за допомогою знака $=$ ($\equiv$ або $\leftrightarrow$): записують $A = B$ ($A \equiv B$ чи $A \leftrightarrow B$).

Відношення рівносильності на множині формул є відношенням еквівалентності, тому часто це відношення називають *еквівалент-ністю*.

Наведемо приклади пар рівносильних формул:

$(A \rightarrow B) = ((\neg A) \vee B)$,

$(\neg(A \vee B)) = ((\neg A) \wedge (\neg B))$, $(\neg(A \wedge B)) = ((\neg A) \vee (\neg B))$,

$(A \wedge (B \vee C)) = ((A \wedge B) \vee (A \wedge C))$,

$(A \vee (B \wedge C)) = ((A \vee B) \wedge (A \vee C))$

тощо.

Ці рівносильності та подібні до них легко перевірити, обчисливши таблиці істинності відповідних функцій для лівих і правих частин і порівнявши ці таблиці.

Цей простий метод можна застосовувати для перевірки рівносильності чи нерівносильності будь-яких формул A і B довільної складності. Тому на перший погляд може здатися, що проблема встановлення рівносильності чи нерівносильності формул алгебри висловлень є розв'язаною і до того ж найпростішим чином і, отже, подальші дослідження в цьому напрямі непотрібні.

Наведемо лише два міркування, які демонструють, що перше враження обманливе.

Перше міркування пов'язане з тим, що коли кількість пропози-ційних змінних у досліджуваних формулах є значною, то застосування зазначеного простого методу може стати практично нездійсненним. Адже навіть для 30 змінних потрібно випробувати більше ніж 10^9 наборів значень змінних для кожної формули. Це тільки кількість кроків загальної процедури, а ще слід урахувати трудомісткість обчислення значення функції істинності на кожному з наборів.

Друге міркування (і це міркування, певно, важливіше), в алгебрі висловлень у більшості випадків цікавляться не рівносильністю двох заданих формул, а рівносильністю нескінченної множини пар формул. Потрібні твердження, згідно з якими всі формули певного типу рівносильні відповідно формулам певного іншого типу. Якщо множини формул обох цих типів нескінченні, то подібні твердження, очевидно, не можна встановити жодним методом, що спирається на побудову таблиць істинності. Для цього потрібні загальні міркування й інші методи.

Зокрема, однією з основних проблем алгебри висловлень є проблема опису класу всіх тавтологій (тотожно істинних формул), яка називається *проблемою розв'язності*. Простішим її варіантом є такий: зазначити правило перевірки тотожної істинності певної формули за допомогою скінченної кількості дій.

Проблема розв'язності посідає важливе місце в математичній логіці. До неї зводиться багато різних задач математичної логіки, наприклад, обговорювана вище проблема перевірки рівносильності заданих формул A і B.

Легко довести таку теорему.

Теорема 3.1. Формули алгебри висловлень A і B рівносильні тоді й тільки тоді, коли формула $((A \to B) \land (B \to A))$ є тавтологією.

Щоб скоротити запис формул, подібних до формули з наведеної теореми, до сигнатури алгебри висловлювань уводять додаткову операцію, що позначається ~ й означається так: $(A \sim B)$ — скорочений запис формули $((A \to B) \land (B \to A))$. Операцію ~ називають ***еквівалентністю*** (*рівнозначністю*).

Отже, останню теорему можна сформулювати так: формули A і B рівносильні тоді й тільки тоді, коли формула $(A \sim B)$ тотожно істинна.

Разом із відношенням рівносильності на множині формул, яке є, як зазначалося, еквівалентністю, розглядають також деякі інші відношення, корисні для логіки та її застосувань. Серед таких виділимо, зокрема, відношення логічного слідування на множині формул алгебри висловлень.

Нехай $A(p_1, p_2, \ldots, p_n)$ і $B(p_1, p_2, \ldots, p_n)$ — формули алгебри висловлень. Будемо говорити, що формула $B(p_1, p_2, \ldots, p_n)$ є ***логічним слідуванням*** (або ***логічним наслідком***) формули $A(p_1, p_2, \ldots, p_n)$, якщо B набуває значення 1 для всіх тих наборів значень пропозиційних змінних, для яких формула A істинна (тобто набуває значення 1); позначатимемо це $A \Rightarrow B$. Іноді в цьому випадку кажуть, що формула A *сильніша*, ніж формула B, або ж B *слабша*, ніж A. Це означає, що множина наборів значень змінних, для яких істинна формула A, є підмножиною множини наборів значень змінних, для яких істинна формула B, що є логічним слідуванням формули A.

Приклад 3.1. Формула $B(x, y, z) = (x \lor z)$ є логічним слідуванням формули $A(x, y, z) = ((x \lor y) \land z)$; це випливає з відповідних таблиць істинності (табл. 3.4).

Дві формули A та B рівносильні тоді й тільки тоді, коли кожна з них є логічним слідуванням іншої, тобто $A = B$ тоді й тільки тоді, коли $A \Rightarrow B$ та $B \Rightarrow A$.

Таблиця 3.4

x y z	A	B
0 0 0	0	0
0 0 1	0	1
0 1 0	0	0
0 1 1	1	1
1 0 0	0	1
1 0 1	1	1
1 1 0	0	1
1 1 1	1	1

З означення випливає також, що будь-яка формула є логічним слідуванням тотожно хибної формули, а тотожно істинна формула (тавтологія) є логічним слідуванням довільної формули.

Проблему перевірки того, чи є формула B логічним слідуванням іншої формули A, також можна звести до проблеми розв'язності для певної формули алгебри висловлень.

Теорема 3.2. Формула $B(p_1, p_2, \ldots, p_n)$ є логічним слідуванням формули $A(p_1, p_2, \ldots, p_n)$ тоді й тільки тоді, коли тотожно істинна формула $(A \to B)$.

Зауважимо, що символам операцій алгебри висловлень відповідають у звичайній мові такі вирази:

$\land$ — "і", "а", "але", "хоч", "незважаючи на…";

$\lor$ — "або", "чи", "хоч одне з…";

$\neg$ — "не", "неправильно, що…";

$\to$ — "якщо…, то…", "…імплікує…", "з… випливає…";

~ — "…тоді й тільки тоді, коли…", "…якщо й тільки якщо…", "…еквівалентне (рівносильне)…".

У математичній мові імплікацію $p \to q$ трактують так: твердження p — достатня умова для q, а твердження q — необхідна умова для твердження p. Вираз $p \to q$ інтерпретують і як "q тоді, коли p" або "p тільки тоді, коли q". Мотивації та об'рунтування такої інтерпретації та вибору наведених у табл. 3.2 означень для логічних операцій можна знайти, наприклад, в [15; 26; 35]. В імплікації $p \to q$ p називають *антецедентом*, або *посилкою*, а q — *консеквентом*, або *висновком*.

Часто для спрощення формул алгебри висловлень за рахунок зменшення числа дужок установлюють такі правила:

1) зовнішні дужки в записі формул випускають;

2) запроваджують такий *порядок виконання* (ранг, або пріоритет) для логічних операцій: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\sim$, який зменшується зліва направо.

Відзначимо, що алгебра висловлень, або, як її іноді називають, *логіка висловлень* є надто бідною теорією для опису логічного апарату математичних міркувань. Типи логічних міркувань, основаних на тотожно істинних формулах алгебри висловлень, далеко не вичерпують логічних законів, використовуваних у математиці, не кажучи вже про логічні міркування в інших науках.

Множина висловлень $M = \{A_1(p_1, p_2, \ldots, p_n), A_2(p_1, p_2, \ldots, p_n), \ldots, A_k(p_1, p_2, \ldots, p_n)\}$ називається ***несуперечною (сумісною)***, якщо існує такий набір значень для $p_1, p_2, \ldots, p_n$, на якому кон'юнкція $A_1 \wedge A_2 \wedge \ldots \wedge A_k$ набуває значення 1 (про висловлення $A_1, A_2, \ldots, A_k$ тоді кажуть, що вони ***сумісні*** між собою). В іншому разі, тобто якщо на всіх наборах $p_1, p_2, \ldots, p_n$ кон'юнкція $A_1 \wedge A_2 \wedge \ldots \wedge A_k$ набуває значення 0, кажуть, що висловлення $A_1, A_2, \ldots, A_k$ ***несумісні*** або, що множина M висловлень ***суперечна***.

При побудові алгебри висловлень вихідними об'єктами були елементарні висловлення, що мали певне значення істинності: 1 або 0. Нові об'єкти — складені висловлення, що також мали певне значення істинності, — будувались за чітко визначеними правилами утворення формул. При цьому значення істинності або хибності складеного висловлення визначають за таблицями істинності відповідних операцій алгебри висловлень. Означені згодом поняття рівносильності і логічного слідування формул введено змістовно, тобто з використанням значень істинності формул залежно від значень їх змінних. Така побудова логічного числення або теорії називається *змістовно-алгоритмічною*, або *табличною*.

Іншим методом побудови логічного числення є описаний вище *формально-аксіоматичний метод*. Саме за його допомогою побудовано так зване числення висловлень.

3.1. Задачі і вправи

1. Чи є наведений вираз висловленням? Якщо так, то визначити, яким саме — істинним чи хибним:

(а) число 1001 кратне 7;

(б) Женева — столиця Швеції;

(в) жодне дійсне число x не задовольняє нерівність $x^2 < 0$;

(г) для будь-яких цілих чисел n і m існує таке ціле число k, що $n < k$ і $k < m$;

(д) неправда, що $2 \cdot 2 = 5$;

(е) ця задача нескладна;

(є) трикутник ABC — правильний;

(ж) "Учітесь, читайте, і чужому научайтесь, й свого не цурайтесь" (*Т. Шевченко*);

(з) викладач процитував: "Учітесь, читайте, і чужому научайтесь, й свого не цурайтесь" (*Т. Шевченко*);

(и) кількість слів у цьому реченні дорівнює семи;

(і) хай живе кібернетика!

(ї) те, що стверджується в цьому реченні, — неправда (*парадокс Епіменіда*).

2. Визначити істинність або хибність складеного висловлення, враховуючи значення елементарних висловлень, з яких воно складається:

(а) Київ розташований на берегах Дніпра, а Дніпро впадає в Чорне море;

(б) число 1001 кратне 17, але не кратне 7;

(в) $-2 > -3$, але $(-2)^2 < (-3)^2$;

(г) якщо $3 < 2$, то $3^2 < 2^2$;

(д) 56 кратне 14 тоді й тільки тоді, коли 56 кратне 2 і 96 кратне 7;

(е) 56 кратне 28 тоді й тільки тоді, коли 56 кратне 8 і 96 кратне 7;

(є) неправильно, що принаймні одне з чисел 1, 21, 101 або 171 є простим.

3. Записати у вигляді формули алгебри висловлень наведене твердження, позначаючи прості висловлення буквами, і визначити його істинність або хибність:

(а) трикутник є рівностороннім тільки тоді, коли він рівнобедрений;

(б) для того щоб трикутник був рівнобедреним, достатньо, щоб він був рівностороннім;

(в) чотирикутник є квадратом тоді, коли він є прямокутником;

(г) для того щоб чотирикутник був квадратом, необхідно, щоб він був прямокутником.

4. Знайти значення формули A:

(а) $A = (a \sim b) \wedge ((\neg c \rightarrow b) \vee (\neg a \wedge c)) \vee (a \wedge \neg b)$ при $a = b = 1$, $c = 0$;

(б) $A = (b \rightarrow a) \wedge (\neg c \vee \neg a) \sim (\neg b \rightarrow a) \wedge (d \vee (\neg b \rightarrow c))$ при $a = c = 0$, $b = d = 1$.

5. Скласти таблицю істинності для формули алгебри висловлень:

(а) $a \rightarrow \neg(b \wedge c) \sim (a \wedge \neg c)$;

(б) $(\neg a \vee b) \sim (\neg b \wedge c)$;

(в) $(a \rightarrow b) \rightarrow ((b \rightarrow c) \rightarrow (a \rightarrow c))$;

(г) $(a \sim \neg b) \wedge (a \wedge b \rightarrow \neg a \wedge \neg b)$.

6. Показати, що дана формула алгебри висловлень є тотожно істинною:

(а) $(a \to b) \to (a \wedge \neg b \to b)$;
(б) $(a \to b) \to \big((b \to c) \to (a \to c)\big)$;
(в) $(a \to \neg a) \to \neg a$;
(г) $(a \to b) \vee (b \to a)$;
(д) $(\neg a \to a) \to a$;
(е) $\big((a \to b) \to a\big) \to a$.

7. Переконатися в тому, що дана формула алгебри висловлень є суперечністю:

(а) $\neg b \wedge a \wedge (a \to b)$;
(б) $a \vee b \sim \neg a \wedge (b \to \neg b)$.

8. Визначити, якою є наведена формула алгебри висловлень: 1) тавтологією; 2) суперечністю; 3) нейтральною:

(а) $a \wedge c \vee b \wedge d \to (a \vee b) \wedge (c \vee d)$;
(б) $(a \wedge c \to b \wedge d) \to (a \to b) \wedge (c \to d)$;
(в) $(a \vee b \to c) \to (a \to c) \vee (b \to c)$;
(г) $\big((a \sim b) \to (c \sim d)\big) \to (a \vee c \to b \vee d)$;
(д) $(a \to b) \wedge (c \to d) \to (a \wedge c \to b \wedge d)$;
(е) $(a \to b) \wedge a \wedge b \sim (a \to b) \wedge a$.

9. Довести або спростувати таке твердження:

(а) якщо формула A алгебри висловлень — тавтологія, то $\neg A$ — суперечність;

(б) з двох формул A, $\neg A$ алгебри висловлень принаймні одна — тотожно істинна;

(в) якщо A і B — тавтології, то $A \wedge B$ — теж тавтологія. Чи правильне обернене твердження?

(г) якщо A і B — тавтології, то $A \vee B$ — тавтологія. Чи правильне обернене твердження?

(д) якщо A і B — тавтології, то $A \to B$ — тавтологія. Чи правильне обернене твердження?

(е) якщо формула $A \sim B$ — тавтологія, то A і B — тавтології.

10. Довести або спростувати таке твердження:

(а) якщо A — виконувана формула алгебри висловлень, то $\neg A$ є невиконуваною;

(б) з двох формул A, $\neg A$ алгебри висловлень хоч одна є виконуваною;

(в) якщо A і B — виконувані формули, то $A \wedge B$ — виконувана. Чи правильне обернене твердження?

(г) якщо A і B — виконувані формули, то $A \vee B$ — виконувана. Чи правильне обернене твердження?

(д) якщо $A \to B$ — виконувана формула, то A і B виконувані;

(е) якщо $A \sim B$ — виконувана формула, то A і B виконувані.

11. Довести, що формула алгебри висловлень, яка із символів логічних операцій містить лише $\wedge$, $\vee$, $\to$, $\sim$, є виконуваною.

12. Довести рівносильність:

(а) $(a \wedge b \to c) \to (a \to c) = \neg a \vee b \vee c$;
(б) $(a \to b) \vee (a \to c) \to \big(a \to (b \to c)\big) = \neg a \vee \neg b \vee c$.

13. Довести або спростувати:

(а) $\neg(a \sim \neg b) = a \wedge b \vee \neg a \wedge \neg b$;
(б) $a \vee c \to b \vee d = (a \to b) \wedge (c \to d)$;
(в) $a \to (a \sim b) = a \to b$;
(г) $a \vee b \vee \neg c \to a = (b \to a) \wedge (\neg a \to c)$.

14. Нехай a і b — якісь висловлення. Визначити, чи є логічно еквівалентними (рівносильними) такі пари тверджень:

(а) "Якщо a, то b" і "Якщо неправильно, що a, то неправильно, що b";

(б) "b — необхідна умова для a" і "b тільки тоді, коли a";

(в) "b — необхідна умова для a" і "$\neg b$ — достатня умова для $\neg a$";

(г) "Неправильно, що a і b" і "Неправильно, що a, і неправильно, що b";

(д) "Неправильно, що a або b" і "Неправильно, що a, і неправильно, що b";

(е) "Неправильно, що a тоді й тільки тоді, коли b" і "a тоді й тільки тоді, коли $\neg b$".

15. Довести, що формули алгебри висловлень A і B рівносильні тоді й тільки тоді, коли формула $(A \sim B)$ є тавтологією.

16. Перевірити, чи має місце така властивість для логічних операцій:

(а) комутативність імплікації;
(б) асоціативність імплікації;
(в) дистрибутивність імплікації щодо кон'юнкції;
(г) дистрибутивність кон'юнкції щодо імплікації;
(д) дистрибутивність імплікації щодо диз'юнкції;
(е) дистрибутивність диз'юнкції щодо імплікації.

17. Відомо, що формула $A \to \neg A$ — тавтологія. Що можна сказати про формулу A?

18. Довести, що формула B алгебри висловлень є логічним слідуванням формули A:

(а) $A = (a \to c) \wedge (b \to c)$, $B = a \wedge b \to c$;
(б) $A = (a \to b) \wedge (b \to c)$, $B = a \to c$ (транзитивна властивість імплікації);
(в) $A = (a \sim b) \wedge (b \sim c)$, $B = a \sim c$ (транзитивна властивість еквівалентності).

19. Відомо, що формула B — логічне слідування формули A, а B — невиконувана формула алгебри висловлень. Що можна сказати про формулу A?

20. Відомо, що B — виконувана формула алгебри висловлень і B сильніша, ніж A. Довести, що формула A теж виконувана.

21. Чи є несуперечною множина висловлень?

(а) $\{\neg(a \to b), b \to a\}$; (г) $\{a \to \neg a, \neg a \to a\}$;
(б) $\{a \wedge \neg b, \neg b \to \neg a\}$; (д) $\{a \to b, \neg a, \neg b\}$;
(в) $\{a \to b, a \to \neg b\}$; (е) $\{a \sim \neg b, \neg a \to \neg c, a \to c, c \to \neg b\}$.

22. Довести, що коли формула $A_1(p_1, \ldots, p_n) \to \big(A_2(p_1, \ldots, p_n) \to \to (\ldots \to A_m(p_1, \ldots, p_n)\ldots)\big)$ невиконувана, то множина формул $\{A_1(p_1, \ldots, p_n), \ldots, A_m(p_1, \ldots, p_n)\}$ суперечна. Чи правильне обернене твердження?

23. Дано три множини M_1, M_2 і M_3. Перевірити, чи є несуперечною множина таких висловлень:

1) якщо $a \in M_1$, то неправильно, що $a \in M_2$;
2) неправильно, що $a \in M_2$ тільки тоді, коли $b \in M_3$;
3) якщо $a \notin M_2$, то $b \in M_3$;
4) неправильно, що $a \notin M_1$ і $b \notin M_3$.

24. Інспектор з'ясовує обставини справи у трьох свідків А, Б і В. А каже інспекторові, щоб він не вірив свідченням Б; Б наполягає на тому, що свідчення В неправдиві, а В твердить, що ані А, ані Б не говорять правди. Хто з трьох свідків каже правду?

25. Троє обвинувачених А, Б, В дають такі свідчення: А — "Б винен, а В — ні"; Б — "Якщо А винен, то і В теж винен"; а В — "Я не винен, але хоч один із двох інших винен":

(а) вважаючи, що Б і В говорять правду, встановити, хто саме винен;
(б) якщо всі троє невинні, то хто з них сказав правду, а хто — неправду?

26. Перевірити, чи є несуперечною така множина висловлень:

1) якщо Андрій не купить продукти, то Ірина не запросить гостей;
2) Тетяна приїде тільки тоді, коли Ірина запросить гостей;
3) Андрій принесе сир, якщо він купить продукти;
4) якщо приїде Тетяна, то Андрій не принесе сир;
5) Ірина не запросить гостей, і Андрій не принесе сир.

3.3. Числення висловлень

Числення висловлень (***ЧВ***) згідно з поданою в розд. 3.1 схемою означають таким чином:

1. ***Алфавіт*** ЧВ складається зі змінних висловлень (пропозиційних змінних): $a, b, c, d, \ldots, x, y, z$ (можливо, з індексами), знаків логічних операцій $\vee$, $\wedge$, $\neg$, $\to$ і круглих дужок (та).

2. Поняття ***формули*** означають аналогічно алгебрі висловлень:

а) усі пропозиційні змінні є формулами;

б) якщо A і B — формули, то вирази (або слова) $(A \wedge B)$, $(A \vee B)$, $(\neg A)$, $(A \to B)$ — також формули;

в) інших формул, ніж побудовані за правилами а) і б), немає.

Відзначимо важливу властивість цього означення. Можна довести, що існує формальна процедура, яка, будучи застосована до будь-якого слова в даному алфавіті, за скінченну кількість кроків установить, чи є дане слово формулою, чи ні. Більше того, за записом формули ця процедура дасть її повний синтаксичний розбір, тобто опис послідовності кроків побудови формули за означеними вище правилами. Зауважимо також, що для зменшення кількості дужок у формулах зазвичай випускають зовнішні дужки формул і замість $(\neg A)$ записують $\neg A$;

3. ***Аксіомами*** $A1$–$A10$ числення висловлень є такі формули:

$A1.$ $a \to (b \to a)$
$A2.$ $(a \to b) \to \big((a \to (b \to c)) \to (a \to c)\big)$
$A3.$ $(a \wedge b) \to a$
$A4.$ $(a \wedge b) \to b$
$A5.$ $(a \to b) \to \big((a \to c) \to (a \to (b \wedge c))\big)$
$A6.$ $a \to (a \vee b)$
$A7.$ $b \to (a \vee b)$
$A8.$ $(a \to c) \to \big((b \to c) \to ((a \vee b) \to c)\big)$
$A9.$ $(a \to b) \to ((a \to \neg b) \to \neg a)$
$A10.$ $\neg\neg a \to a$

4. ***Правилами виведення*** є:

1) ***правило підстановки***. Якщо A — вивідна формула, яка містить літеру p (позначимо $A(p)$), то вивідною є й формула $A(B)$, що одержується з A заміною всіх входжень літери p на довільну формулу B; це записують

$$\frac{A(p)}{A(B)};$$

2) ***правило висновку*** (***modus ponens***). Якщо A і $A \to B$ — вивідні формули, то вивідною є й формула B; це записують

$$\frac{A, A \to B}{B}.$$

У поданому описі ЧВ аксіоми є формулами числення (згідно з означенням формули); формули, використовувані у правилах виведення (A, $A \to B$ тощо), є *метаформулами*, або так званими *схемами формул*.

Схема формул — це вираз метамови для позначення нескінченної множини всіх тих формул числення, які отримують після заміни змінних метамови (*метазмінних*) цієї схеми певними формулами числення.

Наприклад, якщо у схемі формул $A \to B$ замінити A на p, а B — на $p \wedge q$, то з даної схеми отримаємо формулу числення $p \to (p \wedge q)$; якщо ж метазмінну A замінити на $p \to q$, а B — на $(p \wedge q) \to p$, то дістанемо формулу $(p \to q) \to ((p \wedge q) \to p)$ тощо.

Використання схем формул можна поширити й на всі аксіоми. Наприклад, якщо в системі аксіом пропозиційні змінні a, b, c замінити метазмінними A, B, C, то отримаємо десять схем аксіом, що задають десять нескінченних множин аксіом. Таким чином приходимо до іншого способу побудови числення висловлень: із нескінченною множиною аксіом (що задається скінченним числом схем аксіом), але без правила підстановки, оскільки воно неявно міститься в поясненні (*інтерпретації*) схем аксіом.

Перший спосіб є більш послідовно конструктивним: усі його засоби явно зафіксовані та скінченні; при оперуванні з численням за допомогою ЕОМ (наприклад, для автоматизації доведень теорем і побудови виведень для заданих формул) він виглядає природнішим.

З іншого боку, другий спосіб більш відповідає математичній традиції тлумачення (інтерпретації) формул: наприклад, алгебрична тотожність $(a + b)^2 = a^2 + 2ab + b^2$ або відомі логарифмічні чи тригонометричні тотожності тлумачать саме як схеми тотожностей, а не як конкретні тотожності, що справджуються лише для конкретних літер. Правило підстановки при цьому мається на увазі і присутнє неявно. Утім, перехід від одного способу побудови числення до іншого є нескладним.

Аксіоми числення висловлень разом із правилами виведення повністю визначають поняття довідної (вивідної) формули в ЧВ, або теореми ЧВ.

Вивідними формулами, або ***теоремами***, числення висловлень є ті й тільки ті формули, які можна вивести з аксіом за допомогою означених правил виведення.

Розглянемо приклади виведення теорем ЧВ.

Приклад 3.2. Доведемо, що формула $a \to a$ є теоремою ЧВ:

1) підставивши в аксіому $A2$ змінну a замість змінної c і $b \to a$ замість b, матимемо вивідну формулу

$$(a \to (b \to a)) \to ((a \to ((b \to a) \to a)) \to (a \to a));$$

2) підформула $a \to (b \to a)$ — це аксіома $A1$. Тоді з вивідних формул $a \to (b \to a)$ і $(a \to (b \to a)) \to ((a \to ((b \to a) \to a)) \to (a \to a))$ за правилом висновку виводимо формулу $(a \to ((b \to a) \to a)) \to (a \to a)$);

3) замінимо в аксіомі $A1$ b на $(b \to a)$, отримаємо $a \to ((b \to a) \to a)$;

4) знову, застосувавши правило висновку до двох останніх формул, матимемо, що формула $a \to a$ вивідна.

Для зручності будемо застосовувати таку форму запису виведення формул у ЧВ:

а) послідовність формул виведення писатимемо у стовпчик, нумеруючи їх у порядку слідування $F1$, $F2$, …;

б) поряд із кожною формулою після двокрапки писатимемо пояснення, що об´рунтовує законність її появи у виведенні.

Правило підстановки $\dfrac{A(p)}{A(B)}$ записуватимемо у вигляді $S_B^p A = A(B)$ або $S(p)(B)A$, а правило висновку $\dfrac{A,\ A \to B}{B}$ — у вигляді $\mathrm{MP}(A, A \to B) = B$.

Тоді останнє виведення набере такого вигляду:

$F1$: $S^{b,\ c}_{b \to a,\ a} A2 = (a \to (b \to a)) \to ((a \to ((b \to a) \to a)) \to (a \to a))$

$F2$: $\mathrm{MP}(A1, F1) = ((a \to ((b \to a) \to a)) \to (a \to a))$

$F3$: $S^{b}_{b \to a} A1 = (a \to ((b \to a) \to a))$

$F4$: $\mathrm{MP}(F3, F2) = (a \to a)$.

Отже, ми довели таку метатеорему ЧВ: $\vdash (a \to a)$.

Важливим та зручним у численні висловлень є означене вище правило виведення формули з певної множини заданих формул, яке дає змогу значно скорочувати подальші виведення.

Наведемо приклади виведення деяких формул із заданих множин формул.

Приклади 3.3. 1. $a \vdash (b \to a)$

$F1$: a

$F2$: $\mathrm{MP}(F1, A1) = b \to a$

2. $a, b, a \to (b \to c) \vdash c$

$F1$: a

$F2$: b

$F3$: $a \to (b \to c)$

$F4$: $\mathrm{MP}(F1, F3) = b \to c$

$F5$: $\mathrm{MP}(F2, F4) = c$.

3. $a, b \vdash (a \wedge b)$

$F1$: $S^{b,\,c}_{a,\,b}A5 = ((a \to a) \to ((a \to b) \to (a \to (a \wedge b))))$

$F2$: $(a \to a)$ (див. приклад 3.2)

$F3$: $\mathrm{MP}(F2, F1) = ((a \to b) \to (a \to (a \wedge b)))$

$F4$: b

$F5$: $S^{a,\,b}_{b,\,a}A1 = (b \to (a \to b))$

$F6$: $\mathrm{MP}(F4, F5) = a \to b$

$F7$: $\mathrm{MP}(F6, F3) = (a \to (a \wedge b))$

$F8$: a

$F9$: $\mathrm{MP}(F8, F7) = (a \wedge b)$.

Безпосередньо з означення поняття вивідності випливають такі твердження для довільної множини формул Г.

Лема 3.1. Якщо $\vdash B$, то $\Gamma \vdash B$.

Лема 3.2. Якщо $A_1, A_2, \ldots, A_n \vdash B$, то $A_1, A_2, \ldots, A_n, \Gamma \vdash B$.

Лема 3.3. Якщо $\Gamma, A \vdash B$ і $\vdash A$, то $\Gamma \vdash B$.

Лема 3.4. Якщо $\Gamma \vdash A$ і $A \vdash B$, то $\Gamma \vdash B$ (транзитивність відношення вивідності).

Будь-яку доведену в численні вивідність вигляду $\Gamma \vdash A$, де Г — множина формул, A — довільна формула, можна розглядати як правило виведення — , яке можна додати до заданої множини правил.

Наприклад, доведену у прикладі 3.3.1 вивідність $a \vdash b \to a$ разом із правилом підстановки можна розглядати як правило $\dfrac{A}{B \to A}$, що можна інтерпретувати так: "Якщо формула A вивідна, то вивідною є й формула $B \to A$, де B — довільна формула". Це правило надалі можна використовувати для побудови нових виведень. Такі правила називатимемо ***похідними***. За допомогою додаткових похідних правил дістаємо можливість скоротити виведення формул, не виконуючи повного виведення. За скороченим виведенням завжди можна побудувати повне, замінюючи кожну формулу, яка є результатом застосування похідного правила виведення, на повне її виведення. Такою, наприклад, є формула $F2$ у прикладі 3.3.3. Інакше кажучи, похідні правила — це "будівельні блоки" при побудові виведень формул числення висловень, кожен з яких замінює кілька кроків звичайного виведення.

Ефективним засобом одержання багатьох важливих і корисних похідних правил виведення є так звана *метатеорема дедукції* (***МТД***); зокрема, саму МТД можна розглядати як таке похідне правило.

Теорема 3.3 (***метатеорема дедукції***). Якщо $\Gamma, A \vdash B$, то $\Gamma \vdash A \to B$, де Г — множина формул (можливо, порожня), A і B — формули.

Доведення. Зауважимо, що доведення метатеорем, на відміну від теорем числення, проводять змістовно, як звичайне математичне доведення.

Будемо виходити з того, що задані аксіоми є схемами аксіом, тобто не будемо користуватися правилом підстановки.

Нехай $\Gamma, A \vdash B$. Тоді існує таке виведення $F_1, F_2, \ldots, F_n$ із Γ, A, що $F_n = B$. Доведемо за індукцією, що для будь-якого $k \le n$ $\Gamma \vdash A \to F_k$.

Розглянемо спочатку випадок $k = 1$, тобто формулу F_1. Як перша формула виведення вона або має бути аксіомою, або міститися в Г, або збігатися з A.

Зі схеми аксіоми $A1$ випливає, що $F_1 \to (A \to F_1)$ — вивідна формула. Якщо F_1 — аксіома або міститься в Г, то за правилом висновку формула $A \to F_1$ є вивідною з Г. Якщо ж $F_1 = A$, то з прикладу 3.2 маємо, що $A \to A$, тобто $A \to F_1$ — вивідна формула. Отже, у будь-якому разі отримаємо $\Gamma \vdash A \to F_1$.

Відтак припустімо, що $\Gamma \vdash A \to F_i$ для довільного $i < k$, і розглянемо формулу F_k. Можливі такі ситуації:

а) F_k — аксіома;

б) F_k міститься в Г;

в) $F_k = A$;

г) F_k є вивідною з деяких попередніх формул F_j і F_l за правилом висновку; тоді формула F_l повинна мати вигляд $F_j \to F_k$.

У випадках а), б) і в) доведення твердження $\Gamma \vdash A \to F_k$ здійснюється аналогічно доведенню для F_1 (у випадках а) та б) — за допомогою схеми аксіоми $A1$; у випадку в) — за допомогою результату прикладу 3.2).

У випадку г) за індуктивним припущенням маємо $\Gamma \vdash A \to F_j$ і $\Gamma \vdash A \to F_l$, де F_l — це $F_j \to F_k$, тобто $\Gamma \vdash A \to (F_j \to F_k)$. Підставимо в схему аксіоми $A2$ A замість a, F_j замість b і F_k замість c. Дістанемо $(A \to F_j) \to ((A \to (F_j \to F_k)) \to (A \to F_k))$. Двічі застосувавши до останньої вивідної формули правило висновку, отримаємо $\Gamma \vdash A \to F_k$. Залишилося покласти $k = n$.

Розглянемо декілька *застосувань* метатеореми дедукції.

1. Поширеним методом у математичних доведеннях є *метод доведення від супротивного*: припустивши, що деяке a є правильним (істинним твердженням), доводимо, що, по-перше, з a можна вивести b, а по-друге, що з a можна вивести $\neg b$, що неможливо; отже, твердження a хибне, тобто істинним твердженням є $\neg a$.

У термінах числення висловлень цей метод можна сформулювати так: "Якщо $\Gamma, a \vdash b$ і $\Gamma, a \vdash \neg b$, то $\Gamma \vdash \neg a$".

Доведемо справедливість цього правила в численні висловлень.

Справді, за теоремою дедукції, якщо $\Gamma, a \vdash b$ та $\Gamma, a \vdash \neg b$, то $\Gamma \vdash a \to b$ і $\Gamma \vdash a \to \neg b$:

$F1$: $a \to b$

$F2$: $a \to \neg b$

$F3$: $\mathrm{MP}(F1, A9) = (a \to \neg b) \to \neg a$

$F4$: $\mathrm{MP}(F2, F3) = \neg a$

Доведене твердження (метатеорему) часто називають *правилом уведення заперечення* й записують у вигляді

$$\frac{\Gamma, A \vdash B;\ \Gamma, A \vdash \neg B}{\Gamma \vdash \neg A}.$$

Крім того, неважко переконатись у справедливості для числення висловлень такого твердження, або метатеореми, яку можна вважати оберненою до метатеореми дедукції (*ОМТД*): "якщо $\Gamma \vdash A \to B$, то $\Gamma, A \vdash B$".

Послідовно маємо:

$F1$: A

$F2$: $A \to B$

$F3$: $\mathrm{MP}(F1, F2) = B$

2. Доведемо тепер *закон виключення третього* $\vdash a \vee \neg a$:

$F1$: $S^{a,\,b}_{a,\,\neg a}\,A6 = a \to (a \vee \neg a)$

$F2$: $S^{a}_{\neg(a \vee \neg a)}(a \to a) = (\neg(a \vee \neg a)) \to (\neg(a \vee \neg a))$ (див. приклад 3.2).

Із формул $F1$ і $F2$ та леми 3.2 маємо (за ОМТД):

$F3$: $\neg(a \vee \neg a),\ a \vdash a \vee \neg a$

$F4$: $\neg(a \vee \neg a),\ a \vdash \neg(a \vee \neg a)$

За доведеним правилом уведення заперечення зі співвідношень $F3$ та $F4$ отримаємо

$F5$: $\neg(a \vee \neg a) \vdash \neg a$

Аналогічно використаємо аксіому $A7$:

$F6$: $S^{a,\,b}_{a,\,\neg a}\,A7 = \neg a \to (a \vee \neg a)$

$F7$: $\neg(a \vee \neg a),\ \neg a \vdash a \vee \neg a$

$F8$: $\neg(a \vee \neg a),\ \neg a \vdash \neg(a \vee \neg a)$

З $F7$ і $F8$ маємо

$F9$: $\neg(a \vee \neg a) \vdash \neg\neg a$

За правилом уведення заперечення з $F5$ і $F9$ дістанемо:

$F10$: $\neg\neg(a \vee \neg a)$

$F11$: $S^{a}_{a \vee \neg a}A10 = \neg\neg(a \vee \neg a) \to (a \vee \neg a)$

$F12$: $\mathrm{MP}(F10, F11) = a \vee \neg a$, тобто $\vdash a \vee \neg a$.

Існують й *інші числення висловлень*, тобто числення з іншими системами аксіом і правилами виведення.

Наприклад, розглянемо числення висловлень ЧВ1, яке використовує тільки логічні операції $\neg$ і $\to$ і має таку систему аксіом:

$S1$. $a \to (b \to a)$

$S2$. $(a \to (b \to c)) \to ((a \to b) \to (a \to c))$

$S3$. $(\neg a \to \neg b) \to ((\neg a \to b) \to a)$.

Правила виведення в новому численні ті самі, що й у старому — правила підстановки та висновку. Якщо в системі аксіом першого числення замінити підформули $(a \vee b)$ на $(\neg a \to b)$, а підформули $(a \wedge b)$ — на $\neg(a \to \neg b)$, то має місце така теорема.

Теорема 3.4. Обидва наведені числення висловлень ЧВ та ЧВ1 рівносильні в тому розумінні, що множини формул, вивідних у кожному з цих числень (множини їх теорем), збігаються між собою.

Доведення теореми полягає в тому, що доводять вивідність усіх аксіом першого числення з аксіом другого, і навпаки (з урахуванням зауважень стосовно $\vee$ та $\wedge$).

Яке з двох числень краще? Однозначної відповіді на це запитання немає. Друге числення компактне та наочне, тому компактнішими є й доведення різних його властивостей. Водночас у багатшому першому численні виведення різноманітних формул, узагалі кажучи, коротші.

Можливі й інші числення, рівносильні двом наведеним [22, 35].

3.2. Задачі і вправи

1. Чи є формулою числення висловлень запропонований вираз? Відповідь обґрунтувати за індуктивним означенням формули:

(а) $((a \vee ((\neg b) \to c)) \to (\neg a))$;

(б) $(((a \wedge (\neg b)) \vee c) \vee (a \to b))$;

(в) $(((\neg a) \to b) \vee \neg((\neg b) \wedge c))$;

(г) $(((a \wedge b) \wedge c) \to ((b \vee c) \vee (\neg d)))$.

2. Перевірити, чи правильно виконано підстановку:

(а) $S^{a}_{b \to a} A1 = (b \to a) \to (b \to (b \to a))$;

(б) $S^{c}_{b \to c} A2 = (a \to b) \to ((a \to c) \to (a \to c))$;

(в) $S^{a}_{b \to a} A1 = (b \to a) \to (b \to a)$;

(г) $S^{\neg a}_{b \to a} A9 = (a \to b) \to ((a \to \neg b) \to (b \to a))$.

Якщо є помилка, зазначити, у чому саме вона полягає.

3. Із множини припущень Г вивести формулу F:

(а) $\Gamma = \{a \to b, a \to c\}$, $F = a \to (b \wedge c)$ (*правило композиції*);

(б) $\Gamma = \{a, b\}$, $F = a \wedge b$ (*правило введення кон'юнкції*);

(в) $\Gamma = \{a \to b, a \to \neg b\}$, $F = \neg a$ (*правило введення заперечення*);

(г) $\Gamma = \{a \to b, b \to c\}$, $F = a \to c$ (*правило ланцюгового висновку, або правило силогізму*).

4. Позначимо через Г довільну скінченну (можливо, порожню) послідовність формул числення висловлень, через A, B, C — довільні формули числення висловлень. Безпосередньо з означення вивідності довести такі властивості символу "$\vdash$":

(а) $C \vdash C$;

(б) якщо $\Gamma, A, B \vdash C$ і $\vdash B$, то $\Gamma, A \vdash C$;

(в) якщо $\Gamma \vdash C$, то $\Gamma, A \vdash C$;

(г) якщо $B, B, \Gamma \vdash C$, то $B, \Gamma \vdash C$;

(д) якщо $\Gamma_1, A, B, \Gamma_2 \vdash C$, то $\Gamma_1, B, A, \Gamma_2 \vdash C$;

(е) якщо $\Gamma_1 \vdash A$ і $\Gamma_2, A \vdash C$, то $\Gamma_1, \Gamma_2 \vdash C$.

Поясніть звичайними словами зміст тверджень (а)–(е).

5. Довести формально теорему:

(а) $(a \wedge a) \to a$;

(б) $a \to (a \wedge a)$;

(в) $(a \vee a) \to a$;

(г) $a \to (a \vee a)$;

(д) $(a \to b) \to (a \to (a \wedge b))$;

(е) $(b \to (a \vee c)) \to ((a \vee b) \to (a \vee c))$;

(є) $(a \to b) \to ((b \vee a) \to b)$;

(ж) $(a \wedge b) \to (a \vee b)$;

(з) $(a \vee b) \to (b \vee a)$;

(и) $((a \wedge b) \to (b \wedge a)$;

(і) $a \to ((a \vee b) \wedge a)$;

(ї) $(a \vee (a \wedge b)) \to a$.

6. Застосовуючи метатеорему дедукції, довести теорему:

(а) $((a \to b) \to a) \to ((a \to b) \to b)$;

(б) $(a \to c) \to (a \to (b \vee c))$;

(в) $(\neg a \to b) \to (\neg b \to a)$;

(г) $(a \to \neg b) \to (b \to \neg a)$;

(д) $(\neg b \to \neg a) \to (a \to b)$;

(е) $a \to \neg\neg a$;

(є) $((a \to b) \to (a \to c)) \to ((a \to (b \to c)))$;

(ж) $(a \to (b \to c)) \to ((a \to b) \to (a \to c))$.

7. Використовуючи метатеорему дедукції та властивості символу "$\vdash$" (див. 3.2.4), довести похідне правило виведення:

(а) якщо $\Gamma, A \vdash C$ і $\Gamma, B \vdash C$, то $\Gamma, A \vee B \vdash C$ (*правило вилучення диз'юнкції*);

(б) якщо $\Gamma, A \vdash B$ і $\Gamma, A \vdash C$, то $\Gamma, A \vdash B \wedge C$ (*правило композиції*);

(в) $\Gamma, A \to B \vdash \neg B \to \neg A$ (*правило контрапозиції*).

8. Обґрунтувати таке похідне правило виведення:

(а) $A \to B, \neg B \vdash \neg A$ (*modus tollens*);

(б) $A, \neg A \vdash B$;

(в) $A \wedge \neg A \vdash B$;

(г) $A \vee B, \neg A \vdash B$.

9. Використовуючи правило введення заперечення, властивості символу вивідності та метатеорему дедукції, довести теорему:

(а) $(\neg a \vee \neg b) \to \neg(a \wedge b)$;

(б) $\neg(a \wedge b) \to (\neg a \vee \neg b)$;

(в) $\neg(a \vee b) \to (\neg a \wedge \neg b)$;

(г) $(\neg a \wedge \neg b) \to \neg(a \vee b)$.

10. Використовуючи теорему $(a \wedge \neg a) \to b$, довести, що формула $\neg(a \wedge \neg a)$ є теоремою ЧВ (*закон суперечності*).

3.4. Числення висловень і алгебра висловлень. Основні проблеми числення висловлень

Довільну формулу F числення висловлень можна змістовно інтерпретувати як складене висловлення, істинність або хибність якого залежить від істинності елементарних висловлень, що входять до нього. Таким чином, кожній формулі F числення висловлень аналогічно тому, як це було зроблено в алгебрі висловлень, можна поставити у відповідність функцію істинності f.

Виникає питання, як пов'язане таке змістовне "істиннісне" тлумачення (інтерпретація) формул числення висловлень з їх формальною вивідністю.

Теорема 3.5. Будь-яка теорема числення висловлень ЧВ є тотожно істинним висловленням.

Доведення. Тотожну істинність усіх аксіом легко перевірити безпосередньою побудовою відповідних таблиць істинності для кожної з них (рекомендуємо зробити це самостійно).

Відтак доведемо, що обидва правила виведення ЧВ перетворюють тотожно істинні формули в тотожно істинні.

Якщо $A(p_1, p_2, \ldots, p_n)$ — тотожно істинна формула, то для довільного набору значень $a_1, a_2, \ldots, a_n$ її пропозиційних змінних висловлення $A(a_1, a_2, \ldots, a_n)$ істинне. Отже, тотожно істинною буде й будь-яка формула $S_{B_1, B_2, \ldots, B_n}^{p_1, p_2, \ldots, p_n} A$, отримана з формули A шляхом підстановки замість пропозиційних змінних $p_1, p_2, \ldots, p_n$ довільних формул $B_1, B_2, \ldots, B_n$, бо останні можуть набувати лише значень 0 або 1.

Доведемо, що коли формули A та $A \to B$ тотожно істинні, то формула B, яку дістаємо з них за правилом висновку, також тотожно істинна. Припустімо супротивне: нехай B не є тотожно істинною формулою, тобто існує набір значень її змінних, на якому вона набуває значення 0. Підставимо цей набір у формулу $A \to B$. Оскільки A — тавтологія, дістанемо вираз $1 \to 0$, значенням якого є 0. Останнє суперечить припущенню про тотожну істинність формули $A \to B$.

Отже, ми переконалися в тому, що всі аксіоми числення висловлень ЧВ є тотожно істинними формулами алгебри висловлень, а застосування обох правил виведення (підстановки і висновку) до тотожно істинних

формул знову дає тотожно істинні формули. Тому всі вивідні формули (теореми) ЧВ є тотожно істинними формулами алгебри висловлень.

Справедливою є й обернена теорема, яку подаємо без доведення.

Теорема 3.6. Будь-яка тотожно істинна формула алгебри висловлень є теоремою числення висловлень.

Дві останні теореми дають змогу розв'язати три важливі проблеми числення висловлень: несуперечності, повноти і розв'язності. Розглянемо їх послідовно.

1. ***Проблема несуперечності.*** Для кожної формальної теорії кардинальними є питання несуперечності. Справді, така теорія будується послідовним приєднанням нових теорем, які можна формально вивести з аксіом за допомогою правил виведення. Отже, немає жодної гарантії, що в цьому процесі ми не дійдемо до суперечності. Інакше кажучи, виникає питання, чи при поступовому нагромадженні теорем формальної теорії не трапиться так, що одна з теорем суперечитиме іншим. Саме так виникає проблема несуперечності числення.

Числення ***несуперечне***, якщо неможливо одночасно вивести з аксіом числення як формулу A, так і її заперечення $\neg A$.

Наслідок 3.1. Числення висловлень ЧВ — несуперечна формальна теорія.

Справді, якщо формула A вивідна в ЧВ, то формула $\neg A$ не може бути вивідною, бо за теоремою 3.5 формула A є тотожно істинною в алгебрі висловлень, а формула $\neg A$ — тотожно хибною. Отже, $\neg A$ не може бути теоремою ЧВ.

2. ***Проблема повноти.*** Інша проблема, що виникає при дослідженні різних числень висловлень: чи будь-яка тотожно істинна формула алгебри висловлень є вивідною в заданому численні? Це питання й являє собою *проблему повноти* для числення висловлень.

Сенс такої постановки питання полягає в тому, що при побудові числення потрібно знати, чи достатньо його аксіом і правил виведення для того, щоб можна було вивести будь-яку формулу, тотожно істинну у змістовному розумінні.

Наслідок 3.2. Числення висловлень ЧВ — повне.

Це твердження безпосередньо випливає з теореми 3.6.

У математичній логіці існує й інше поняття повноти системи аксіом (або числення), що ґрунтується на неможливості доповнити систему аксіом будь-якою формулою, яку не можна вивести з даних аксіом.

3. ***Проблема розв'язності. Розв'язувальним методом*** для формальної теорії T називають метод, за допомогою якого для довільної формули A з T можна за скінченне число кроків визначити, чи є A теоремою, чи ні.

Числення T називають ***розв'язним***, якщо для нього існує розв'язувальний метод, а інакше формальна теорія T нерозв'язна.

Наслідок 3.3. Числення висловлень ЧВ — розв'язна теорія.

Доведення. Нехай A — довільна формула ЧВ. Побудуємо для неї таблицю істинності та розглянемо її останній стовпчик. Якщо він містить лише одиниці, то A — тотожно істинна формула, і за теоремою 3.6 вона є теоремою ЧВ. Інакше (якщо останній стовпчик таблиці містить хоча б один нуль), A — не тавтологія, а тому A не є теоремою. Зрозуміло, що всі ці дії можна виконати за скінченне число кроків.

Нарешті, розглянемо ще одну важливу проблему для числень.

Система аксіом числення називається ***незалежною***, якщо жодну з аксіом цієї системи не можна вивести з інших аксіом системи.

Зрозуміло, що аксіому, яку можна вивести з інших, можна виключити з системи аксіом, і при цьому множина теорем теорії залишиться тією самою (тобто отримаємо рівносильне числення). Отже, залежна система аксіом у певному розумінні не така досконала, як незалежна, бо вона містить зайві аксіоми.

Можна довести, що системи аксіом числень ЧВ та ЧВ1 незалежні.

Існують й інші формальні теорії, що означають і досліджують у математичній логіці: *числення предикатів*, різноманітні *числення* (*теорії*) *першого порядку*, *числення з рівностями*, *формальна арифметика* тощо. У наступних розділах розглянемо основні ідеї та принципи побудови однієї з таких теорій — числення предикатів.

3.5. Поняття предиката

Розглянуте в попередніх розділах числення висловлень є важливою і невід'ємною складовою всіх числень математичної логіки. Однак воно занадто бідне для опису й аналізу найпростіших логічних міркувань науки та практики.

Одна з причин цього полягає в тому, що у численні висловлень будь-яке просте висловлення розглядають як вихідний об'єкт дослідження, неподільне ціле, без частин і внутрішньої структури, яке має лише одну властивість — бути або істинним, або хибним.

Щоб побудувати систему правил, яка давала б змогу проводити логічні міркування для виведення нетривіальних правильних висновків з урахуванням будови і змісту простих висловлень, запропоновано формальну теорію, що дістала назву *числення предикатів*.

Теорія предикатів починається з аналізу граматичної будови простих висловлень і ґрунтується на такому висновку: прості висловлення виражають той факт, що деякі об'єкти (або окремий об'єкт) мають певні властивості, або, що вони перебувають між собою в певному відношенні.

Наприклад, в істинному висловленні "3 — просте число" підмет "3" — це об'єкт, а присудок "просте число" виражає певну його властивість.

У латинській граматиці присудок називається *предикатом*; звідси цей термін і ввійшов у математичну логіку. Головною для логіки предикатів є саме друга складова речення-висловлення — присудок-властивість. Її фіксують, а значення об'єкта пропонують змінювати так, щоб кожен раз отримувати змістовні речення, тобто висловлення.

Наприклад, замінюючи в наведеному вище висловленні 3 на числа 1, 5, 9 або 12, матимемо відповідно такі висловлення: "1 — просте число", "5 — просте число", "9 — просте число", "12 — просте число", з яких друге істинне, а решта — хибні висловлення.

Таким чином, можна розглянути вираз "x — просте число", який не є висловленням, а є так званою *пропозиційною* (*висловлювальною*) *формою*, тобто формою (або формуляром), після підстановки в яку замість параметра (змінної) x об'єктів (значень) із певної множини M дістаємо висловлення.

Аналогічно можна трактувати, наприклад, пропозиційні форми "a — українець", "b і c — однокурсники", "c важче, ніж d", або "точка x лежить між точками y і z". У перші дві з них можна підставляти замість параметрів a, b і c прізвища конкретних людей, у третю — замість c і d назви будь-яких об'єктів (предметів), що мають вагу. Для четвертої множиною M значень змінних x, y і z може бути множина точок певної прямої.

Перша з цих пропозиційних форм задає, як і в наведеній раніше формі, певну властивість для об'єкта a. Інші три форми описують деякі відношення між об'єктами.

Розглянувши конкретні приклади й коротко зупинившись на мотивації та змістовній інтерпретації подальших понять, перейдемо до формальних математичних означень.

n*-місним предикатом** $P(x_1, x_2, \ldots, x_n)$ на множині M називається довільна функція типу $M^n \to B$, де $B = \{0, 1\}$ — булевий (двійковий) алфавіт. Множина M називається ***предметною областю, або *універсальною множиною*, а $x_1, x_2, \ldots, x_n$ — ***предметними змінними***, або *термами*, предиката P.

Множина елементів $(a_1, a_2, \ldots, a_n) \in M^n$ таких, що $P(a_1, a_2, \ldots, a_n) = 1$, називається ***областю істинності*** (або ***характеристичною множиною***) предиката P.

Якщо $P(a_1, a_2, \ldots, a_n) = 1$, то згідно з логічною інтерпретацією будемо говорити, що предикат P є *істинним* на $(a_1, a_2, \ldots, a_n)$. Інакше казатимемо, що предикат P — *хибний* на $(a_1, a_2, \ldots, a_n)$.

Коли універсальна множина скінченна й порівняно невелика, логічну функцію, яку визначає предикат, можна задавати *таблицею*.

Узагалі кажучи, можна означити так званий *багатосортний предикат* як функцію типу $M_1 \times M_2 \times \ldots \times M_n \to B$, дозволивши різним його аргументам набувати значення з різних множин. Іноді це доцільно, однак частіше в логіці предикатів використовують наведене раніше означення.

Неважко зрозуміти, що пропозиційна форма — це один зі способів задання предиката.

Для $n = 1$ предикат $P(x)$ називається ***одномісним***, або *унарним*, для $n = 2$ $P(x, y)$ — ***двомісним***, або *бінарним*, для $n = 3$ $P(x, y, z)$ — ***тримісним***, або *тернарним*.

Коли в n-арному предикаті $P(x_1, x_2, \ldots, x_n)$ зафіксувати деякі m змінних (тобто надати їм певних значень із множини M), то отримаємо $(n - m)$-місний предикат на множині M. Тому можна вважати висловлення нульмісними предикатами, які утворено з багатомісних предикатів підстановкою замість усіх їх параметрів певних значень із предметної області. Отже, висловлення можна розглядати як окремий випадок предиката.

Для довільної множини M і довільного n існує взаємно однозначна відповідність між сукупністю всіх n-місних предикатів на M і множиною всіх n-арних відношень на M. А саме, будь-якому предикату $P(x_1, x_2, \ldots, x_n)$ відповідає відношення R таке, що $(a_1, a_2, \ldots, a_n) \in R$ тоді й тільки тоді, коли $P(a_1, a_2, \ldots, a_n) = 1$. Очевидно, що R — область істинності предиката P.

Крім того, за будь-якою відповідністю C між множинами A і B (тобто $C \subseteq A \times B$) можна побудувати бінарний двосортний предикат $P(x, y)$

у такий спосіб: $P(a, b) = 1$ тоді й тільки тоді, коли $(a, b) \in C$ для $a \in A$ і $b \in B$.

Зокрема, будь-якій функціональній відповідності або функції $f: M^n \to M$ можна поставити у відповідність $(n + 1)$-місний предикат P на M такий, що $P(a_1, a_2, \ldots, a_n, a_{n+1}) = 1$ тоді й тільки тоді, коли $f(a_1, a_2, \ldots, a_n) = a_{n+1}$.

Отже, такі фундаментальні математичні поняття, як відповідність (зокрема, функція), відношення, висловлення можна розглядати як окремі випадки більш загального поняття предиката.

3.6. Логіка предикатів

Як з елементарних висловлень за допомогою логічних операцій можна утворювати складені висловлення, так і, використовуючи прості (елементарні) предикати та логічні зв'язки (операції), можна будувати складені предикати, або *предикатні формули*.

Зазвичай основні логічні операції $\wedge$, $\vee$, $\neg$, $\to$, $\sim$ означають для предикатів, що задані на одній і тій самій предметній області M і залежать від тих самих змінних.

Нехай $P(x_1, x_2, \ldots, x_n)$ і $Q(x_1, x_2, \ldots, x_n)$ — n-місні предикати на множині M.

Кон'юнкцією $P(x_1, x_2, \ldots, x_n) \wedge Q(x_1, x_2, \ldots, x_n)$ предикатів $P(x_1, x_2, \ldots, x_n)$ і $Q(x_1, x_2, \ldots, x_n)$ називають предикат $R(x_1, x_2, \ldots, x_n)$, що набуває значення 1 на тих і тільки тих наборах значень термів, на яких обидва предикати $P(x_1, x_2, \ldots, x_n)$ і $Q(x_1, x_2, \ldots, x_n)$ набувають значення 1.

Область істинності предиката $R(x_1, x_2, \ldots, x_n) = P(x_1, x_2, \ldots, x_n) \wedge$ $\wedge\, Q(x_1, x_2, \ldots, x_n)$ збігається з теоретико-множинним перетином областей істинності предикатів $P(x_1, x_2, \ldots, x_n)$ і $Q(x_1, x_2, \ldots, x_n)$.

Диз'юнкцією $P(x_1, x_2, \ldots, x_n) \vee Q(x_1, x_2, \ldots, x_n)$ предикатів $P(x_1, x_2, \ldots, x_n)$ і $Q(x_1, x_2, \ldots, x_n)$ називають предикат $T(x_1, x_2, \ldots, x_n)$, що набуває значення 1 на тих і тільки тих наборах значень термів, на яких або предикат $P(x_1, x_2, \ldots, x_n)$, або предикат $Q(x_1, x_2, \ldots, x_n)$ набуває значення 1.

Областю істинності предиката $T(x_1, x_2, \ldots, x_n)$ є об'єднання областей істинності предикатів $P(x_1, x_2, \ldots, x_n)$ і $Q(x_1, x_2, \ldots, x_n)$.

Запереченням $\neg P(x_1, x_2, \ldots, x_n)$ предиката $P(x_1, x_2, \ldots, x_n)$ називають предикат $S(x_1, x_2, \ldots, x_n)$, що набуває значення 1 на тих і лише тих значеннях термів, на яких предикат $P(x_1, x_2, \ldots, x_n)$ набуває значення 0.

Областю істинності предиката $S(x_1, x_2, \ldots, x_n) = \neg P(x_1, x_2, \ldots, x_n)$ є доповнення (до множини M^n) області істинності предиката $P(x_1, x_2, \ldots, x_n)$.

Аналогічно вводять й інші логічні операції: $\to$, $\sim$ тощо. Для кожної з них існує певна теоретико-множинна формула, що задає область істинності складеного предиката, і операндами якої є області істинності відповідних предикатів-аргументів. Неважко узагальнити означення всіх уведених операцій для предикатів $P(x_1, x_2, \ldots, x_n)$ і $Q(y_1, y_2, \ldots, y_m)$, що залежать від різних змінних і мають різну місність.

Знаючи, як виконуються окремі операції, можна утворювати вирази або формули, операндами яких є предикати. Наприклад, розглянемо формулу $P_1(x) \vee (\neg P_3(x, z) \to P_2(y, x, z))$, що задає деякий предикат $Q(x, y, z)$. Значення предиката Q неважко обчислити для будь-якого набору значень його термів x, y, z, виходячи зі значень предикатів P_1, P_2, P_3 на цьому наборі.

3.3. Задачі і вправи

1. Побудувати таблицю одномісного предиката:

(а) "кратне 2 і кратне 3"; (б) "кратне 2 або кратне 3" на множині натуральних чисел $\{1, 2, \ldots, 20\}$.

2. Побудувати таблицю двомісного предиката, заданого нерівністю $|x + y| > 1$, де числа x та y належать предметній області $M = \{-2, -1, 0, 1, 2\}$.

3. Виразити множину істинності предиката через множини істинності відповідних елементарних предикатів:

(а) $P(x) \to Q(x)$;

(б) $P(x) \sim Q(x)$;

(в) $(P(x) \vee \neg Q(x)) \wedge R(x) \vee \neg R(x) \to \neg P(x)$;

(г) $\neg(P(x) \vee Q(x)) \sim (\neg P(x) \vee \neg Q(x))$.

4. Нехай на універсальній множині людей предикат Б(a, b) означає : "a — батько b", М(a, b) означає: "a — мати b", Ч(a) означає: "a — чоловік". Записати через Б, М, Ч такий предикат:

(а) a і b — брати;

(б) a — небіж b;

(в) a — дідусь b;

(г) a — онук, а b — онучка c.

3.7. Квантори

Додатково в логіці предикатів використовують дві спеціальні операції, які називають ***кванторами***. За допомогою цих операцій, по-перше, можна перетворювати пропозиційні форми у висловлення, а по-друге, теорія предикатів стає значно гнучкішою, глибшою і багатшою, ніж теорія висловлень. Саме тому логіку предикатів іноді називають *теорією квантифікації*.

Найпопулярнішими й найуживанішими виразами в математиці є фрази або формулювання типу "для всіх" та "існує". Вони входять до більшості проміжних математичних міркувань або доведень чи остаточних тверджень, висновків, лем або теорем, наприклад: "*Для всіх* дійсних чисел x виконується рівність $\sin^2 x + \cos^2 x = 1$", "Для заданих натуральних a і b завжди *існує* натуральне число d, яке ділить числа a і b", "*Для всіх* натуральних n справедливе твердження: якщо n ділиться на 6 і на 15, то n ділиться на 30" — тощо.

Поняття, що відповідає словам "для всіх", лежить в основі означення квантора загальності.

Нехай $P(x)$ — предикат на множині M. Тоді ***квантор загальності*** — це операція, що ставить у відповідність предикату $P(x)$ висловлення "для всіх x з M $P(x)$ істинно". Для позначення цієї операції використовують знак $\forall$, який і називають квантором загальності. Останнє висловлення в математичній логіці записують $\forall x\ P(x)$ (читають: "для всіх x P від x").

Існує й інший квантор, що є в певному розумінні двоїстий до квантора загальності і називається ***квантором існування***. Позначається він знаком $\exists$. Якщо $Q(x)$ — предикат на множині M, то висловлення "існує у множині M елемент x такий, що $Q(x)$ істинно" записують у вигляді $\exists x\ Q(x)$ (читають: "існує такий x, що Q від x" або "є такий x, що Q від x").

Походження обраних позначень пояснюється тим, що символ $\forall$ — це перевернута велика перша літера німецького слова "alle" або англійського слова "all", що перекладаються "усі", а символ $\exists$ відповідає першій літері слів "existieren" (нім.) або "exist" (англ.) — існу-вати.

Вираз $\forall x$ читають також як "усі x", "для кожного x", "для довільного x", "для будь-якого x", а вираз $\exists x$ — як "деякий x", "для деякого x", "знайдеться такий x" тощо.

Окрім уведених символічних позначень кванторів, використовують й інші позначення. Так, замість $\forall x$ іноді пишуть $\forall(x)$, (x) або $\wedge x$, а замість $\exists x$ відповідно — $\exists(x)$, $(\mathrm{E}x)$ або $\vee x$.

Приклад 3.4. Розглянемо два бінарні предикати на множині натуральних чисел: предикат "x менше y" і предикат "x ділить y". Перший з них запишемо в традиційній формі $x < y$, а другий — у вигляді $x \mid y$. Тоді неважко переконатися, що висловлення $\forall x\ \exists y\ (x < y)$ і $\forall x\ \exists y\ (x \mid y)$ істинні, а висловлення $\exists y\ \forall x\ (x < y)$ і $\exists y\ \forall x\ (x \mid y)$ — хибні. Істинними є, наприклад, висловлення $\forall x\ (((2 \mid x) \wedge (3 \mid x)) \to (6 \mid x))$, $\forall x\ (0 < x^2 - x + 1)$, $\forall x\ ((x < 1) \to (x < 2))$, $\exists x\ ((x \mid 1) \wedge (\neg(1 < x)))$, а висловлення $\forall x\ ((3 \mid x) \to (6 \mid x))$, $\exists x\ (x^2 < 0)$, $\forall x\ (2 \mid x)$ хибні.

Важливу роль у логіці предикатів відіграє поняття *області дії квантора*, під якою розумітимемо той вираз, до якого відноситься цей квантор. Область дії квантора позначають за допомогою дужок. Ліву дужку, що відповідає початку області дії, записують безпосередньо після кванторної змінної квантора, а відповідна їй права дужка означає закінчення області дії цього квантора. Там, де це не спричиняє невизначеності, дужки можна випускати й замість $\forall x\ (P(x))$ або $\exists x\ (P(x))$ писати відповідно $\forall x\ P(x)$ або $\exists x\ P(x)$.

Приклад 3.5. В усіх наведених кванторних виразах область дії квантора підкреслено: $\exists x\ (\underline{(3 \mid x) \to (6 \mid x)})$, $\forall x\ (\underline{(x^2 < 9) \to (x < 3)})$, $\exists x\ (\underline{3 \mid x}) \to (6 \mid x)$, $\forall x\ (\underline{x^2 < 9}) \to (x < 3)$.

Перший і третій вирази з останнього прикладу, а також другий і четвертий відрізняються не лише областю дії квантора. Відмінність між ними істотніша, і про це слід сказати окремо.

Розглянемо на універсальній множині R дійсних чисел два вирази $x^2 < 10$ та $\exists x\ (x^2 < 10)$. Перший з них є предикатом, що залежить від змінної x. Замість x у нього можна підставляти різні дійсні значення й отримувати певні висловлення (істинні чи хибні). Та сама предметна змінна x входить у другий вираз інакше. Якщо замість неї підставити будь-яке дійсне значення, дістанемо беззмістовний вираз.

Нехай $P(x)$ — деякий предикат на множині M. Перехід від $P(x)$ до $\forall x\ P(x)$ або $\exists x\ P(x)$ називається ***зв'язуванням*** змінної x. Інші назви — *навішування квантора* на змінну x (або на предикат $P(x)$), або *квантифікація* змінної x.

Змінна x, на яку навішено квантор, називається ***зв'язаною***, в іншому разі змінна x називається ***вільною***.

Зв'язана змінна, як було продемонстровано вище, вже не є змінною в класичному розумінні цього поняття. Вона потрібна лише для ідентифікації терма, від якого залежить відповідна пропозиційна форма. Вираз $\forall x\ P(x)$ не залежить від x і при фіксованих P й M має певне значення. Звідси, зокрема, випливає, що можна здійснювати перейменування зв'язаної змінної (тобто переходити від $\forall x\ P(x)$ до $\forall t\ P(t)$) і воно не змінює значення істинності виразу.

Зауважимо, що розглядувана ситуація не виняткова й доволі часто зустрічається в інших розділах математики. Наприклад, у виразах $\int_a^b f(x)dx$, $\lim_{x \to c} x^2$ та $\sum_{j=k}^{n} d_j$ параметри a, b, c, k і n — це змінні, замість яких можна підставляти певні значення, а параметри x і j — зв'язані змінні, підстановка замість яких будь-яких значень не має жодного сенсу.

Навішувати квантори можна й на багатомісні предикати. Наприклад, застосовуючи квантори $\forall$ і $\exists$ до змінних x і y двомісного предиката $A(x, y)$, отримаємо чотири різні одномісні предикати $\forall x\ A(x, y)$, $\exists x\ A(x, y)$, $\forall y\ A(x, y)$ і $\exists y\ A(x, y)$. У перших двох змінна x зв'язана, а змінна y вільна, а у двох останніх — навпаки.

Вираз $\forall x\ A(x, y)$ (читається "для всіх x A від x та y") є одномісним предикатом $B(y)$. Він істинний для тих і тільки тих $b \in M$, для яких одномісний предикат $A(x, b)$ істинний для всіх x з M.

Приклад 3.6. Розглянемо двомісний предикат $A(x, y)$, визначений на множині $M = \{a_1, a_2, a_3, a_4\}$ за допомогою табл. 3.5.

Таблиця 3.5

$x \backslash y$	a_1	a_2	a_3	a_4
a_1	0	1	1	0
a_2	0	1	1	1
a_3	0	0	1	1
a_4	0	0	1	0

Таблиці істинності для чотирьох відповідних одномісних предикатів, отриманих з $A(x, y)$ навішуванням одного квантора, наведено в табл. 3.6.

Таблиця 3.6

y	$\forall x\, A(x, y)$	y	$\exists x\, A(x, y)$	x	$\forall y\, A(x, y)$	x	$\exists y\, A(x, y)$
a_1	0	a_1	0	a_1	0	a_1	1
a_2	0	a_2	1	a_2	0	a_2	1
a_3	1	a_3	1	a_3	0	a_3	1
a_4	0	a_4	1	a_4	0	a_4	1

В усіх чотирьох випадках до вільної змінної, що залишилася, можна застосувати один із кванторів і, зв'язавши таким чином обидві змінні, перетворити відповідні предикати у висловлення.

Для предиката з останнього прикладу отримаємо такі висловлення:

$$\forall x\, (\forall y\, A(x, y)) = 0, \qquad \forall y\, (\forall x\, A(x, y)) = 0,$$
$$\exists x\, (\exists y\, A(x, y)) = 1, \qquad \exists y\, (\exists x\, A(x, y)) = 1,$$
$$\exists y\, (\forall x\, A(x, y)) = 1, \qquad \exists x\, (\forall y\, A(x, y)) = 0,$$
$$\forall y\, (\exists x\, A(x, y)) = 0, \qquad \forall x\, (\exists y\, A(x, y)) = 1.$$

Неважко переконатися, що висловлення, які містять однакові квантори, рівносильні. Обидва висловлення $\forall x\, (\forall y\, A(x, y))$ і $\forall y\, (\forall x\, A(x, y))$ істинні тоді й тільки тоді, коли предикат $A(x, y)$ набуває значення 1 на всіх кортежах значень (a, b) з M^2. Висловлення $\exists x\, (\exists y\, A(x, y))$ і $\exists y\, (\exists x\, A(x, y))$ істинні тоді й тільки тоді, коли існує принаймні одна така пара (a, b), що $A(a, b) = 1$.

Водночас усі чотири висловлення з різнойменними кванторами є, взагалі кажучи, нерівносильними. Особливо слід наголосити, що суттєвим є порядок слідування різнойменних кванторів.

Висловлення $\forall x\, (\exists y\, A(x, y))$ і $\exists y\, (\forall x\, A(x, y))$, узагалі кажучи, нерівносильні. Наприклад, у термінах табличного задання предиката $A(x, y)$ істинність першого висловлення $\forall x\, (\exists y\, A(x, y))$ означає, що кожен рядок таблиці істинності містить принаймні одну одиницю. А друге висловлення $\exists y\, (\forall x\, A(x, y))$ істинне тоді й лише тоді, коли в таблиці є стовпчик, що складається тільки з одиниць.

Неважко поширити всі наведені вище міркування й висновки на предикати більшої арності. Навішування одного квантора завжди зменшує число вільних змінних і арність предиката на одиницю. Застосування кванторів до всіх змінних предиката перетворює його у висловлення (іноді таку предикатну формулу називають ***замкненою***).

3.4. Задачі і вправи

1. Записати символікою логіки предикатів такі твердження:

(а) існує таке x, що $P(x)$ — хибне;

(б) "існує таке x, що $P(x)$" — хибне;

(в) "для кожного x $P(x)$" — хибне;

(г) $P(x)$ — хибне для кожного x.

2. Зазначити вільні і зв'язані змінні у даному виразі. Для кожної зв'язаної змінної визначити, яким саме квантором її зв'язано:

(а) $P(x) \to \forall y\, (Q(y) \vee \exists z\, P(z) \sim \neg P(y))$;

(б) $\exists x\, (P(y) \to P(x)) \wedge \exists z\, (Q(z) \sim \exists y\, (Q(y) \vee Q(x))) \wedge \forall x\, (P(x) \to \neg(\exists y\, Q(y)))$;

(в) $\forall x\, (x^2 y > 0 \to y > 0)$;

(г) $x > 3 \wedge \forall x\, \forall y\, (xy^2 > 0)$;

(д) $\forall x\, \forall y\, (x < y \to \exists z\, (x < z \wedge z < y))$.

3. У цій вправі предметна область — це множина R дійсних чисел. Визначити, чи є даний вираз висловленням або пропозиційною формою. У першому випадку зазначити, істинним чи хибним є висловлення; у другому навісити квантори так, щоб одержати істинне висловлення:

(а) $\exists y\, (x < y)$;

(б) $\exists y\, \forall x\, (xz = xy)$;

(в) $\forall x\, \forall y\, \exists z\, (xy = z)$;

(г) $\exists x\, \forall y\, \exists z\, (xy = z)$;

(д) $\exists p\, \forall x\, (x^2 + px + q > 0)$;

(е) $\exists y\, \forall x\, (xy > 0) \to \exists z\, (z^2 < 0)$.

4. Нехай предметною областю є множина N натуральних чисел. Проаналізувати область дії кванторів і знайти значення висловлення:

(а) $\forall x\, \forall y\, (\exists z\, (z > x \wedge z < y) \sim x < y)$;

(б) $\forall x\, \forall y\, \exists z\, ((z > x \wedge z < y) \sim x < y)$.

5. Навести приклади тверджень як математичного, так і нематематичного змісту, в яких є кванторні вирази: "для кожного" й "існує", і значення яких змінюється при зміні порядку слідування цих виразів.

6. Порівняти області дії квантора та значення істинності виразів $\forall x\, (x > 3) \to (2 > 3)$ і $\forall x\, ((x > 3) \to (2 > 3))$.

3.8. Формули. Рівносильність формул. Тотожно істинні формули

Наведемо індуктивне означення поняття ***формули*** *логіки предикатів* (*предикатної формули*, або просто *формули*) на предметній області M.

1. Усі предикати $P(x_1, x_2, \ldots, x_n)$ на множині M є формулами. Їх називають ***елементарними***, або ***атомарними***.

2. Якщо A і B — формули, то $(\neg A)$, $(A \wedge B)$, $(A \vee B)$, $(A \to B)$, $(A \sim B)$ — теж формули.

3. Якщо A — формула, а x — вільна змінна в A, то $(\forall x\,(A))$ і $(\exists x\,(A))$ — теж формули.

4. Інших формул, крім утворених за правилами 1–3, немає.

Це означення дає змогу твердити, що всі формули алгебри висловлень є формулами логіки предикатів, оскільки висловлення — це нульмісні предикати. За допомогою означення неважко також переконатися, що вирази $(\forall x\ (\exists y\ (A(x, y)) \to (B(x) \vee (\exists z\ (C(x, z))))))$ і $(\forall x\ (\forall y\ (A(x, y) \wedge B(x)) \to (\exists y\ (C(x, y)))))$ є формулами логіки предикатів, а $(\forall x (A(y) \to (\exists x (B(x)))))$ — не формула, оскільки у виразі $(A(y) \to (\exists x (B(x))))$, який є правильною формулою, змінна x зв'язана, тобто не є вільною змінною, і до неї не можна застосувати квантор $\forall x$.

Для зручності можна запровадити такі домовленості про скорочення кількості дужок у формулах. По-перше, залишимо всі умови скорочення кількості дужок, які було прийнято в алгебрі висловлень, виходячи з пріоритету логічних операцій. По-друге, випускатимемо всі зовнішні дужки. Вважатимемо, що квантори мають більший пріоритет, ніж логічні операції. Випускатимемо також дужки, що позначають область дії квантора, якщо остання є елементарною формулою. Нарешті, не писатимемо дужки між кванторами, що слідують один за одним. При цьому такі кванторні операції виконуються в порядку, зворотному до їх написання (справа наліво).

Нехай $F(x_1, x_2, \ldots, x_n)$ — деяка формула логіки предикатів на множині M. У разі логічної (істиннісної) інтерпретації формули F можливі такі три основні ситуації.

1. Існує набір значень змінних, для якого формула F набуває значення 1 або перетворюється на істинне висловлення. У цьому разі формула F називається ***виконуваною в області*** M.

Якщо для формули F існує область M, у якій вона виконувана, то формула F називається просто ***виконуваною***.

2. Якщо формула F набуває значення 1 (виконувана) для всіх наборів значень з області M, то вона називається ***тотожно істинною в*** M. Формула, тотожно істинна в будь-яких областях M, називається ***тотожно істинною***, або ***логічно загальнозначущою*** (скорочено — ***ЛЗЗ***).

3. Якщо формула F невиконувана в області M, то вона називається ***тотожно хибною в*** M. Формула, невиконувана в усіх областях M, називається ***тотожно хибною***, або ***суперечністю***.

Приклад 3.7. Формула $\exists x\, A(x, y) \to \forall x\, A(x, y)$ виконувана й тотожно істинна в усіх одноелементних областях M. Формула $F(x_1, x_2, \ldots, x_n) \vee \neg F(x_1, x_2, \ldots, x_n)$ тотожно істинна, а формула $F(x_1, x_2, \ldots, x_n) \wedge \neg F(x_1, x_2, \ldots, x_n)$ тотожно хибна. Формули $\forall x\, P(x) \to P(y)$ і $P(y) \to \exists x\, P(x)$ є тотожно істинними.

Формули F_1 і F_2 називають ***рівносильними*** (***еквівалентними***), якщо при всіх можливих підстановках значень замість їх змінних вони набувають однакових значень; позначають $F_1 = F_2$.

Наприклад, усі тотожно істинні (усі тотожно хибні) формули рівносильні між собою. Очевидно також, що коли F_1 і F_2 рівносильні, то формула $F_1 \sim F_2$ тотожно істинна, і навпаки.

Множина всіх тотожно істинних формул логіки предикатів є складовою всіх формальних математичних теорій, тому її дослідження й опис — важлива задача математичної логіки. Значення цієї множини підкреслює й той факт, що їй, як було зазначено вище, належать усі рівносильні співвідношення (тотожності) логіки предикатів.

Як і в логіці висловлень, постають дві проблеми: 1) опис або побудова множини всіх тотожно істинних формул; 2) перевірка тотожної істинності заданої формули логіки предикатів.

Якщо є процедура розв'язання другої проблеми, то на її основі можна сформулювати такий тривіальний алгоритм, що породжує шукану множину T тотожно істинних формул. Послідовно будуємо всі формули, кожну з них за відомою процедурою перевіряємо на тотожну істинність і вносимо до множини T ті, для яких результат перевірки позитивний.

Однак на відміну від логіки висловлень, де така процедура існує та зводиться до обчислення значень даної формули на скінченній множині значень її параметрів, у логіці предикатів області визначення предметних і предикатних змінних формул, узагалі кажучи, нескінченні (зліченні чи навіть незліченні).

Метод обчислення значення формули шляхом підстановки значень замість змінних і послідовного виконання зазначених дій є зручним для встановлення виконуваності заданої формули або доведення нерівносильності певних формул. Для цього достатньо підібрати одну відповідну підстановку. Застосовувати цей метод можна також, коли предметна область M скінченна. Це пов'язано з тим, що для скінченної множини $M = \{a_1, a_2, \ldots, a_n\}$ кванторні формули можна перетворити в рівносильні їм звичайні формули логіки висловлень:

$$\forall x\, P(x) = P(a_1) \wedge P(a_2) \wedge \ldots \wedge P(a_n),$$
$$\exists x\, P(x) = P(a_1) \vee P(a_2) \vee \ldots \vee P(a_n).$$

Замінивши всі квантори за допомогою наведених співвідношень, можна будь-яку формулу логіки предикатів перетворити в рівносильну пропозиційну форму, або формулу логіки висловлень. Істинність останньої на скінченній множині M можна перевірити за допомогою скінченної кількості підстановок і обчислень.

Для доведення ж рівносильності предикатних формул, заданих на нескінченних предметних областях, прямий перебір непридатний, і доводиться застосовувати різні опосередковані методи.

Наприклад, вище за допомогою простих міркувань було доведено рівносильність формул, що описують переставність однойменних кванторів у двомісних предикатах, тобто доведено тотожну істинність формул $\forall x\, \forall y\, A(x, y) \sim \forall y\, \forall x\, A(x, y)$ і $\exists x\, \exists y\, A(x, y) \sim \exists y\, \exists x\, A(x, y)$.

Аналогічними міркуваннями доведемо рівносильність, що описує дистрибутивність квантора $\forall x$ щодо кон'юнкції:

$$\forall x\, (A(x) \wedge B(x)) = \forall x\, A(x) \wedge \forall x\, B(x).$$

Нехай ліва частина цього співвідношення істинна для деяких предикатів A і B. Тоді для будь-якого $a \in M$ істинна кон'юнкція $A(a) \wedge B(a)$. Тому $A(a)$ і $B(a)$ одночасно істинні для довільних a, отже формула $\forall x\, A(x) \wedge \forall x\, B(x)$ істинна. Якщо ж ліва частина хибна, то це означає, що для деякого $a \in M$ хибне або $A(a)$, або $B(a)$. Тому хибне або $\forall x\, A(x)$, або $\forall x\, B(x)$, а отже, хибною є і права частина.

Подібним методом можна довести дистрибутивність квантора $\exists x$ щодо диз'юнкції:

$$\exists x\, (A(x) \vee B(x)) = \exists x\, A(x) \vee \exists x\, B(x).$$

Водночас аналогічні прості міркування дозволяють переконатись, що квантори $\forall x$ і $\exists x$, взагалі кажучи, недистрибутивні щодо диз'юнкції і кон'юнкції. Істинні лише такі імплікації:

$$\forall x\, A(x) \vee \forall x\, B(x) \to \forall x\, (A(x) \vee B(x)),$$
$$\exists x\, (A(x) \wedge B(x)) \to \exists x\, A(x) \wedge \exists x\, B(x).$$

Якщо один із предикатів $A(x)$ або $B(x)$ тотожно істинний, то ліва і права частини першої імплікації водночас істинні. Якщо ж існуватимуть такі значення $a, b \in M$, що $A(a)$ і $B(b)$ хибні, то ліва частина хибна, а права може бути або хибною, або істинною. Для її істинності достатньо, щоб для кожного $a \in M$ істинним був принаймні один із предикатів. Це означає, що знак імплікації $\to$ не можна замінити на знак еквівалентності $\sim$, отже ліва і права частини першої імплікації нерівносильні.

Пропонуємо самостійно проаналізувати другу імплікацію і довести її істинність.

Доведемо ще одне корисне і популярне в логіці й математиці рівносильне співвідношення:

$$\neg(\exists x\, P(x)) = \forall x\, (\neg P(x)).$$

Нехай для деякого предиката P і предметної області M ліва частина істинна. Тоді не існує $a \in M$, для якого $P(a)$ істинне. Отже, для всіх $a \in M$ $P(a)$ хибне, тобто $\neg P(a)$ істинне. Тому права частина істинна. Якщо ж ліва частина хибна, то існує $b \in M$, для якого $P(b)$ істинне, тобто $\neg P(b)$ хибне. Отже, права частина також хибна.

Аналогічно можна довести рівносильність

$$\neg(\forall x\, P(x)) = \exists x\, (\neg P(x)).$$

Наведемо без доведень ще декілька важливих рівносильних співвідношень. Нехай B — предикатна формула, що не містить вільних входжень змінної x. Тоді справджуються такі рівносильності:

$$\forall x\, (A(x) \vee B) = \forall x\, A(x) \vee B, \qquad B \to \forall x\, A(x) = \forall x\, (B \to A(x)),$$
$$\exists x\, (A(x) \vee B) = \exists x\, A(x) \vee B, \qquad B \to \exists x\, A(x) = \exists x\, (B \to A(x)),$$
$$\forall x\, (A(x) \wedge B) = \forall x\, A(x) \wedge B, \qquad \forall x\, A(x) \to B = \exists x\, (A(x) \to B),$$
$$\exists x\, (A(x) \wedge B) = \exists x\, A(x) \wedge B, \qquad \exists x\, A(x) \to B = \forall x\, (A(x) \to B).$$

Ці співвідношення означають, що формулу, яка не містить вільних входжень x, можна виносити за межі області дії квантора, що зв'язує x. З іншого боку, ці самі рівносильності дають змогу включати або вносити відповідну формулу B до області дії квантора за змінною x, від якої B не залежить.

Можливість проведення зазначених рівносильних перетворень для предикатних формул дають змогу означити в логіці предикатів поняття певної *канонічної*, або *нормальної* форми.

Формула, що має вигляд $Q_1x_1Q_2x_2\ldots Q_nx_nF$, де $Q_1, Q_2, \ldots, Q_n$ — квантори, а F — формула, яка не містить кванторів і є областю дії всіх

n кванторів, називається ***випередженою*** (***пренексною***) ***нормальною формулою***, або формулою у *випередженій формі*.

Формула у випередженій формі, рівносильна формулі F, називається ***випередженою*** (***пренексною***) ***формою*** формули F.

Використовуючи останні вісім рівносильних співвідношень і деякі інші, індукцією за числом логічних операцій можна довести, що для кожної формули логіки предикатів існує її випереджена нормальна форма [15; 22; 35].

3.5. Задачі і вправи

1. Довести, що формула логіки предикатів $\forall x\, \exists y\, (P(x) \vee \neg P(y))$ логічно загальнозначуща. Чи є такою формула $\exists y\, \forall x\, (P(x) \vee \neg P(y))$?

2. Довести, що дана формула тотожно істинна:

(а) $\exists x\, (P(x) \wedge Q(x)) \to (\exists x\, P(x) \wedge \exists x\, Q(x))$;

(б) $(\forall x\, P(x) \to \exists x\, Q(x)) \sim \exists x\, (P(x) \to Q(x))$.

3. Довести чи спростувати твердження про те, що запропонована формула тотожно істинна:

(а) $\exists x\, P(x) \wedge \exists x\, (P(x) \to Q(x)) \to \exists x\, Q(x)$;

(б) $\exists x\, P(x) \wedge \forall x\, (P(x) \to Q(x)) \to \exists x\, Q(x)$;

(в) $\forall x\, P(x) \wedge \exists x\, (P(x) \to Q(x)) \to \exists x\, Q(x)$;

(г) $\forall x\, (P(x) \to (Q(x) \to R(x))) \to \forall x\, (P(x) \vee Q(x) \to R(x))$.

4. Довести, що формули A і B логіки предикатів рівносильні:

(а) $A = \exists x\, (P(x) \vee Q(x))$, $B = \exists x\, P(x) \vee \exists x\, Q(x)$;

(б) $A = \neg(\forall x\, P(x))$, $B = \exists x\, (\neg P(x))$.

5. Побудувавши відповідну інтерпретацію (контрприклад), довести, що формули A і B логіки предикатів не є рівносильними:

(а) $A = \exists x\, (P(x) \wedge Q(x))$, $B = \exists x\, P(x) \wedge \exists x\, Q(x)$;

(б) $A = \forall x\, (P(x) \vee Q(x))$, $B = \forall x\, P(x) \vee \forall x\, Q(x)$;

(в) $A = \exists x\, P(x) \to \exists x\, Q(x)$, $B = \exists x\, (P(x) \to Q(x))$.

6. Довести чи спростувати твердження про рівносильність формул A і B логіки предикатів:

(а) $A = \forall x\, (P(x) \to (Q(x) \to R(x)))$, $B = \forall x\, (Q(x) \to (P(x) \to R(x)))$;

(б) $A = \forall x\, (P(x) \to (Q(x) \to R(x)))$, $B = \forall x\, (P(x) \to (R(x) \to Q(x)))$;

(в) $A = \forall y\, \forall x\, (P(x) \to Q(y))$, $B = \forall y\, (\forall x\, P(x) \to Q(y))$;

(г) $A = \forall x\, (P(x) \to Q(y))$, $B = \forall x\, (P(x)) \to Q(y)$;

(д) $A = \exists x\, (P(x) \to Q(y))$, $B = \exists x\, (P(x)) \to Q(y)$.

7. Визначити, які з тверджень 1–5 логічно еквівалентні:

1) неправильно, що всі числа, кратні 4, є точними квадратами;

2) усі числа, кратні 4, не є точними квадратами;

3) не всі числа, кратні 4, є точними квадратами;

4) існує число, кратне 4, яке не є точним квадратом;

5) деякі числа, не кратні 4, не є точними квадратами.

*У наступних вправах **8–12** перевірити, чи випливають на базі логіки предикатів із даних припущень зроблені з них висновки, тобто чи коректні проведені міркування. Якщо міркування правильне, то побудувати відповідний дедуктивний ланцюжок; якщо ж проведене міркування некоректне, то побудувати відповідний контрприклад.*

8. 1) Деякі вписані в коло чотирикутники є прямокутниками;

2) кожний прямокутник є паралелограмом.

Отже, деякі вписані в коло чотирикутники є паралелограмами?

9. 1) Кожне число, кратне 51, кратне 17 і кратне 3;

2) кожне число кратне 3 тільки тоді, коли сума цифр його запису в десятковій системі числення кратна 3;

3) сума цифр запису числа 10712 не кратна 3.

Отже,число 10712 не кратне 51?

10. 1) Кожна злічення множина не є скінченною;

2) кожна незлічення множина є нескінченною;

3) дана множина A — скінченна;

4) жодна скінченна множина не є нескінченною.

Отже, існує множина, яка не є зліченною, і не є незліченною?

11. 1) Кожна нескінченна множина рівнопотужна деякій своїй власній підмножині;

2) жодна скінченна множина не рівнопотужна жодній своїй власній підмножині.

Отже, жодна скінченна множина не є нескінченною і жодна нескінченна множина не є скінченною?

12. 1) Якщо хтось може розв'язати дану задачу, то знайдеться здібний студент, який зробить це;

2) Петренко — здібний студент, але не зміг розв'язати цю задачу.

Отже, ніхто не здатен розв'язати цю задачу?

13. Провести повне доведення твердження: "Множина всіх ірраціональних чисел має потужність не меншу, ніж злічення" — з таких посилок:

1) для кожної нескінченної множини існує її злічення підмножина;

2) множина всіх ірраціональних чисел нескінченна;

3) кожна множина має потужність не меншу, ніж будь-яка її підмножина.

14. Показати, що з множини припущень:

1) V — множина всіх множин;

2) для кожної множини M існує множина T, потужність якої більша, ніж потужність M;

3) якщо M — підмножина множини T, то потужність M не більша, ніж потужність T;

4) кожна множина є підмножиною V

можна вивести суперечність.

Тим самим показати, що поняття множини всіх множин у теорії множин Кантора є суперечним.

15. Показати, що з посилки: "У місті NN цирульник — це той його мешканець, який голить усіх тих і тільки тих мешканців міста NN, хто не голить себе сам", — логічно випливає (на базі логіки предикатів), що в місті NN немає жодного цирульника (*парадокс цирульника*).

16. Замінити формулу $\neg(\forall x\ P(x)) \wedge \exists x\ (\neg Q(x))$ рівносильною формулою логіки предикатів, яка була б у випередженій нормальній формі.

3.9. Числення предикатів. Теорія першого порядку

Числення предикатів, тобто формальну теорію предикатів, будують за вищенаведеною класичною схемою побудови формальних (математичних) теорій.

1. ***Алфавіт*** числення предикатів, тобто множина вихідних символів, складається з ***предметних (індивідних) змінних*** x_1, x_2, ..., ***предметних (індивідних) констант*** a_1, a_2, ..., ***предикатних букв*** $P_1^1, P_2^1, \ldots, P_k^j, \ldots$ і ***функціональних букв*** $f_1^1, f_2^1, \ldots, f_k^j, \ldots$ а також знаків логічних операцій $\vee$, $\wedge$, $\neg$, $\rightarrow$, кванторів $\forall$, $\exists$ і розділових знаків (,) , , (кома).

Верхні індекси предикатних і функціональних букв указують на число аргументів (арність), а нижні використовують для звичайної нумерації букв.

2. Поняття формули означають у два етапи.

Спочатку означають поняття ***терма***:

а) предметні змінні та предметні константи є термами;

б) якщо f^n — функціональна буква, а t_1, t_2, ..., t_n — терми, то $f^n(t_1, t_2, \ldots, t_n)$ — терм;

в) інших термів, крім утворених за правилами а) і б), немає.

Відтак формулюють означення ***формули***.

а) якщо P^n — предикатна буква, а $t_1, t_2, \ldots, t_n$ — терми, то $P^n(t_1, t_2, \ldots, t_n)$ — формула, яка називається ***елементарною***; усі входження предметних змінних у формулу $P^n(t_1, t_2, \ldots, t_n)$ називають ***вільними***;

б) якщо F_1, F_2 — формули, то вирази $(\neg F_1)$, $(F_1 \vee F_2)$, $(F_1 \wedge F_2)$, $(F_1 \rightarrow F_2)$ — теж формули; усі входження змінних, вільні у F_1 і F_2, є вільними й в усіх чотирьох видах формул;

в) якщо $F(x)$ — формула, що містить вільні входження змінної x, то $\forall x\ F(x)$ і $\exists x\ F(x)$ — формули; у цих формулах усі входження змінної x називають ***зв'язаними***, а входження решти змінних у формулу F залишаються вільними;

г) інших формул, крім побудованих за правилами а) – в), немає.

Зауваження. Функціональні символи і терми введено в означення для потенційних потреб різноманітних конкретних прикладних числень предикатів. У прикладних численнях предметна область M є, зазвичай, носієм певної алгебричної системи, тому в численні доцільно мати засоби для опису операцій і відношень, заданих на M. Чисте числення предикатів будують для довільної предметної області; структура цієї області та зв'язки (відношення) між її елементами не беруть до уваги, тому вводити в ньому функціональні букви і терми не обов'язково.

3. ***Аксіоми*** числення предикатів утворюють дві групи:

а) першу групу складають аксіоми довільного числення висловлень (наприклад, можна взяти будь-яку з вищенаведених двох систем $A1$–$A10$ або $S1$–$S3$). Зазвичай ці аксіоми є схемами аксіом;

б) до другої групи належать так звані *предикатні аксіоми*:

$P1$. $\forall x\ F(x) \rightarrow F(y)$,

$P2$. $F(y) \rightarrow \exists x\ F(x)$.

У цих аксіомах $F(x)$ — будь-яка формула, що містить вільні входження x, причому жодне з них не лежить в області дії квантора за y. Формулу $F(y)$ отримуємо з $F(x)$ заміною всіх вільних входжень змінної x на y. Останні зауваження означають, що формула $F(x)$ не може мати, наприклад, вигляд $\exists y\ A(x, y)$, $\forall y\ (A(x) \rightarrow B(y))$ тощо.

4. ***Правилами виведення*** у численні предикатів є такі:

а) правило ***висновку*** — те саме, що й у численні висловлень;

б) правило ***узагальнення*** (*уведення квантора* $\forall$): з $A \rightarrow B(x)$ виводиться $A \rightarrow \forall x\ B(x)$;

в) правило *введення квантора* $\exists$: із $B(x) \rightarrow A$ виводиться $\exists x\ B(x) \rightarrow A$.

В обох останніх правилах формула $B(x)$ містить вільні входження x, а формула A їх не містить.

У нашому численні правила підстановки немає. Отже, з двох можливих методів побудови числення обрано метод зі схемами аксіом. Побудова числення предикатів з правилом підстановки можлива, однак вона є істотно громіздкішою через потребу розрізняти при підстановках вільні та зв'язані входження предметних змінних. Тому частіше в логіці використовують підхід зі схемами аксіом.

Поняття виведення (доведення) формули, теореми, виведення формули з множини гіпотез означають у численні предикатів так само, як це було зроблено в розд. 3.1 для загальної формальної теорії. Справджуються також теореми, подібні до теорем 3.5 і 3.6 числення висловлень.

Теорема 3.7. Будь-яка вивідна формула (теорема) числення предикатів тотожно істинна (логічно загальнозначуща (ЛЗЗ)).

Цю теорему можна довести аналогічно теоремі 3.5. Спочатку слід перевірити, що всі аксіоми є ЛЗЗ формулами. Відтак потрібно довести, що всі правила виведення зберігають властивість ЛЗЗ.

Теорема 3.8. Будь-яка тотожно істинна предикатна формула є вивідною (теоремою) у численні предикатів.

Доведення цієї теореми досить складне, тому тут не наводиться.

З останніх теорем випливає твердження, подібне до твердження теореми 3.1.

Теорема 3.9. Предикатні формули A і B рівносильні ($A = B$) тоді й тільки тоді, коли формула $((A \to B) \wedge (B \to A))$ вивідна в численні предикатів, тобто є ЛЗЗ.

Як і раніше, для скорочення виразу $((A \to B) \wedge (B \to A))$ вводять операцію $\sim$ і записують цей вираз у вигляді $(A \sim B)$. Отже, останню теорему можна переформулювати так: формули A і B рівносильні тоді й тільки тоді, коли формула $(A \sim B)$ вивідна в численні предикатів.

Оскільки, як уже було зазначено вище, виявлення рівносильності формул у логіці предикатів є значно складнішою задачею, ніж у логіці висловлень, то важливе значення останнього твердження полягає в тому, що цю задачу можна звести до пошуку формального виведення для відповідної формули.

Побудоване числення предикатів називають ***численням предикатів першого порядку***, або *теорією першого порядку*. У такій теорії аргументами функцій і предикатів, а також змінними, зв'язаними кванторами, можуть бути лише предметні змінні. У численнях другого та вищих порядків аргументами предикатів можуть бути й предикати, а квантори можуть зв'язувати й предикатні змінні, тобто допустимі, наприклад, вирази вигляду $\forall P\,(P(x))$. Такі числення застосовують значно рідше, тому в математичній логіці їм приділяють менше уваги.

3.6. Задачі і вправи

1. Довести, що формули логіки предикатів $\forall x\,(P(x) \to P(y))$ і $\forall x\,(P(x)) \to P(y)$ рівносильні на будь-якій одноелементній множині $\{a\}$.

2. Довести, що формули з попередньої вправи нерівносильні на множині з трьох елементів, побудувавши відповідну інтерпретацію.

3. Показати, що одна з формул логіки предикатів A або B логічно загальнозначуща (тотожно істинна), а друга — ні:
(а) $A = \forall x\,(P(x)) \to P(y)$, $B = \forall x\,(P(x) \to P(y))$;
(б) $A = \exists x\,(P(x)) \to P(y)$, $B = \exists x\,(P(x) \to P(y))$.

4. Довести, що дана формула логіки предикатів є ЛЗЗ:
(а) $\forall x\, F(x) \to F(y)$; (б) $F(y) \to \exists x\, F(x)$.
У цих виразах $F(x)$ — будь-яка формула, що містить вільні входження x, причому жодне з них не лежить в області дії квантора за y. Формулу $F(y)$ отримуємо з $F(x)$ заміною всіх вільних входжень змінної x на y.

5. Чи є y вільною змінною для x у формулі:
(а) $\exists y\, F(x, y)$; (в) $P(x) \to \forall y\, Q(y)$;
(б) $\forall y\,(F(x, y)) \vee P(x)$; (г) $Q(y) \to \forall z P(x, y, z)$?

6. Чи є змінна x вільною для y у формулі:
(а) $Q(x) \wedge \forall x\,(P(x, y)) \vee P(z, x)$;
(б) $Q(z, y) \vee \exists x\, R(x, y, z)$?

7. Довести, що коли формула $A \to B(x)$ є ЛЗЗ, то $A \to \forall x\, B(x)$ також ЛЗЗ.

8. Довести, що коли формула $B(x) \to A$ є ЛЗЗ, то $\exists x\, B(x) \to A$ також ЛЗЗ.

У задачах **7** і **8** формула $B(x)$ містить вільні входження x, а формула A їх не містить.

9. Навести контрприклади, які спростовують твердження двох попередніх задач, якщо не виконано застережень стосовно характеру входження змінної x у формули A і B.

3.10. Застосування логіки предикатів

Числення предикатів, яке не містить функціональних букв і предметних констант, називається *чистим*. Досі мова йшла переважно саме про чисте числення предикатів. Такі числення містять тільки означені вище так звані ***логічні*** аксіоми (або схеми аксіом).

Прикладні числення (теорії першого порядку) характеризуються тим, що в них до логічних аксіом додають власні ***спеціальні*** аксіоми, у яких означено властивості конкретних предикатних (індивідних) букв і предметних констант із певної предметної області.

Найтиповіші приклади індивідних предикатних букв — предикати = (рівності) і ≤ (порядку), а функціональних букв — знаки арифметичних операцій +, ×, −, / тощо й інших популярних математичних функцій. Як предметні області найчастіше використовують множину N натуральних чисел, множину Z цілих чисел, множину R дійсних чисел, булеан $\beta(A)$ деякої множини A та ін.

Більшість прикладних числень містить ***предикат рівності*** = й аксіоми, що його означають. Наприклад, аксіомами для рівності можуть бути такі:

$E1$. $\forall x\ (x = x)$,

$E2$. $\forall x\ \forall y\ ((x = y) \to (y = x))$,

$E3$. $\forall x\ \forall y\ \forall z\ ((x = y) \to ((y = z) \to (x = z)))$.

Будь-яка теорія, у якій $E1$, $E2$ й $E3$ є аксіомами або теоремами, називається ***теорією*** (або *численням*) ***з рівністю***.

Аналогічно можна ввести три аксіоми, що задають більш загальний предикат — ***предикат еквівалентності*** $E(x, y)$:

$Q1$. $\forall x\ E(x, x)$,

$Q2$. $\forall x\ \forall y\ (E(x, y) \to E(y, x))$,

$Q3$. $\forall x\ \forall y\ \forall z\ ((E(x, y) \wedge E(y, z)) \to E(x, z))$.

Інше прикладне числення — ***теорія часткового порядку***, яка містить три конкретні аксіоми для предиката ≤:

$O1$. $\forall x\ (x \le x)$,

$O2$. $\forall x\ \forall y\ (((x \le y) \wedge (y \le x)) \to (x = y))$,

$O3$. $\forall x\ \forall y\ \forall z\ ((x \le y) \to ((y \le z) \to (x \le z)))$.

Приєднавши до них аксіому

$O4$. $\forall x\ \forall y\ ((x \le y) \vee (y \le x) \vee (x = y))$,

дістанемо *теорію лінійного (досконалого) порядку*.

Ще одна аксіома (*аксіома щільності*)

$O5$. $\forall x\ \forall y\ ((x \le y) \to \exists z\ ((x \le z) \wedge (z \le y)))$

формалізує відношення лінійного порядку в щільних множинах, наприклад у множині раціональних або множині дійсних чисел (див. розд. 1.8).

Найбільш дослідженою на сьогодні формальною теорією, що відіграє визначальну роль для аналізу проблеми об´рунтування засад математики, є так звана ***формальна арифметика***.

У формальній арифметиці використовують три функціональні букви +, ×, $'$. Є також одна предикатна буква — символ бінарного предиката рівності = й одна предметна константа 0.

Дев'ять схем спеціальних аксіом задають основні закони формальної арифметики:

$A1$. $F(0) \wedge \forall x\ (F(x) \to F(x'))\ \to \forall y\ F(y)$ (*принцип індукції*),

$A2$. $\forall x\ \forall y\ ((x' = y') \to (x = y))$,

$A3$. $\forall x\ (\neg(x' = 0))$,

$A4$. $\forall x\ \forall y\ \forall z\ ((x = y) \to ((x = z)\ \to (y = z)))$,

$A5$. $\forall x\ \forall y\ ((x = y) \to (x' = y'))$,

$A6$. $\forall x\ (x + 0 = x)$,

$A7$. $\forall x\ \forall y\ (x + y' = (x + y)')$,

$A8$. $\forall x\ (x \times 0 = 0)$,

$A9$. $\forall x\ \forall y\ (x \times y' = x \times y + x)$.

Зауважимо, що формальна арифметика припускає так звану ***стандартну інтерпретацію***, в якій символ = ототожнюють зі звичайним знаком рівності, 0 — із числом нуль, + і × — із традиційними знаками арифметичних бінарних операцій додавання і множення, а $'$ — з унарною операцією "безпосередньо слідує за". Така інтерпретація відповідає звичній змістовній арифметиці. Кожен терм відповідає якомусь натуральному числу, а формула — твердженню про певну властивість натуральних чисел або числових змінних.

Ретельні дослідження формальної арифметики дали змогу видатному австрійському математику й логіку Курту Геделю і його послідовникам отримати в 30-х роках XX століття фундаментальні результати в галузі реалізації задекларованої на межі XIX та XX століть іншим видатним математиком Давидом Гільбертом програми формального об´рунтування математики. Дві славетні *теореми Геделя про неповноту* знаменували новий етап розвитку математики.

У результаті дослідження різних теорій математики дійшли висновку, що їх об´рунтування можна звести до дослідження систем аксіом для елементарної арифметики, з одного боку, і теорії множин — з іншого. Такими дослідженнями з початку XX століття займалося багато математиків. І лише на початку 30-х років К. Гедель опублікував досить несподіваний на той час і песимістичний результат: жодна скінченна система аксіом для елементарної арифметики не є повною. Точніше, перша теорема Геделя твердить, що будь-яка формальна теорія T, що містить формальну арифметику, неповна, тобто в T існує (і може бути ефективно побудована) така замкнена формула F, що формула $\neg F$ істинна, однак ні F, ні $\neg F$ не є вивідними в T. Друга теорема Геделя про неповноту стверджує, що для довільної несуперечливої формальної теорії T, що включає формальну арифметику, формула, що описує несуперечність T, є невивідною в T. (Тут доречно зауважити, що при доведенні першої з теорем К. Гедель використав метод, подібний до відомого *діагонального методу* Кантора.)

Отже, ні для арифметики та теорії чисел, ні тим більше для багатших математичних теорій не існує адекватних формалізацій. Цей досить сумний, але об'єктивний факт, однак, не заперечує та не знецінює ідеологію формалізму. Формальний підхід залишається основним конструктивним засобом побудови й дослідження математичних теорій. Потенційна неможливість адекватної та повної формалізації теорії означає, що належить або виділяти лише ті фрагменти теорії, які можна формалізувати, й обмежуватися ними, або ж будувати іншу потужнішу формальну теорію (на жаль, знову неповну), котра розширить сферу дії формалізму. Зокрема, використавши метод трансфінітної індукції, який не можна формалізувати у формальній арифметиці, представник гільбертівської школи Герхард Генцен довів несуперечність формальної арифметики й окремих розділів математичного аналізу.

Значення та цінність будь-якої формальної теорії визначається врешті-решт її здатністю описувати певні системи об'єктів і взаємозв'язки між ними. Тому одне з найперших питань до будь-якого числення таке: для опису яких об'єктів придатна ця формальна теорія? У зв'язку з цим вводять поняття інтерпретації для формальної теорії.

Нехай $A = \langle M, \Omega, \pi \rangle$ — алгебрична система. Інтерпретація формальної теорії T складається з множини M і відображення γ, що ставить у відповідність кожній предикатній букві P_k^j певне n-місне відношення $R \in \pi$ на множині M, кожній функціональній букві f_k^j — n-арну операцію $\varphi \in \Omega$ на M, кожній предметній константі — елемент множини M. Постійні терми числення (які не містять предметних змінних) також відображаються в елементи множини M. Носій M алгебричної системи називають ***областю інтерпретації***, а саму алгебричну систему — ***інтерпретацією*** формальної теорії T.

Наприклад, встановлюючи зв'язок між численням висловлень й алгеброю висловлень у розд. 3.4, ми, насправді, здійснили інтерпретацію

теорії ЧВ. Областю інтерпретації M було обрано множину всіх висловлень, а символи (букви алфавіту) $\wedge$, $\vee$, $\neg$, $\to$ було відображено у відповідні операції алгебри висловлень.

У результаті інтерпретації кожна замкнена формула теорії (така, що не містить вільних змінних) перетворюється в певне висловлення стосовно об'єктів алгебричної системи A. Це висловлення може бути або істинним, або хибним. Відкритій формулі (тобто незамкненій) відповідає деяке відношення R на M. При підстановці замість вільних змінних певних предметних констант з M дана формула теж перетворюється у висловлення про те, що ці елементи множини M перебувають у відношенні R.

Відкрита формула F називається ***виконуваною*** *в даній інтерпретації*, якщо існує такий набір предметних констант, підстановка якого у формулу F перетворює її на істинне висловлення.

Формула називається ***істинною в даній інтерпретації***, якщо вона виконувана в разі підстановки довільного набору предметних констант. Формула називається ***хибною в даній інтерпретації***, якщо вона невиконувана. Інтерпретація формальної теорії називається ***правильною***, якщо всі аксіоми теорії є істинними в цій інтерпретації.

Один із методів установлення несуперечності теорії T — побудова її правильної інтерпретації. Зокрема, цим методом було доведено несуперечність геометрії Лобачевського відносно евклідової геометрії, а питання про несуперечність останньої було зведено до проблеми несуперечності арифметики. Метод інтерпретації дає змогу також розв'язувати проблему незалежності системи аксіом.

У цьому розділі, присвяченому елементам математичної логіки, неможливо в достатній мірі висвітлити ці актуальні й цікаві проблеми. Детальніше та глибше з історією та сучасним станом досліджень у галузі математичної логіки й об'рунтуванням засад математики можна (і варто) ознайомитися зі спеціальної літератури [7; 15; 17; 22; 35 та ін.].

Окрім суто формальних побудов у класичному численні предикатів мову так званого ***вузького числення предикатів*** використовують для запису тверджень (властивостей, аксіом, лем, теорем) й означень у різних конкретних розділах математики. Використання символіки логіки предикатів дає змогу досягти більшої строгості та формальності у викладенні математичних результатів, уникнути неоднозначності й багатослівності звичайної мови. Досвід свідчить, що засвоєння методики символічного запису сприяє як полегшенню розуміння змісту досить складних математичних тверджень, так і успішній побудові багатоетапних логічних ланцюжків для розв'язання конкретних задач.

Наприклад, твердження про те, що довільне ціле число a можна розділити з остачею на ціле число b, яке не дорівнює нулю, можна записати так:

$$\forall(a\in Z)\ \forall(b\in Z)\ [(b\neq 0)\rightarrow(\exists(q\in Z)\ \exists(r\in Z)\ (a=b\times q+r)\wedge$$

$$((r=0)\vee((0<r)\wedge(r<|\,b\,|))))].$$

Часто, коли предметна область відома та не змінюється, замість $\forall(a\in Z)$ записують просто $\forall a$. У наведеному виразі всі предикатні букви для позначення відношень $=$, $\neq$, $<$, $\in$ і всі знаки арифметичних і логічних операцій мають звичайний смисл. Записане твердження читається так: "Для цілих a та b, якщо b не дорівнює нулю, існують цілі числа q й r, для яких $a = bq + r$ і r або дорівнює 0, або r більше нуля та менше $|\,b\,|$".

Предикатні формули зручно використовувати для запису означень різних понять. Вище за їх допомогою було означено відношення (предикати) рівності, еквівалентності та порядку. Подібним чином можна записати, наприклад, означення предиката $x\,|\,y$ "x ділить y", або "x — дільник y" на множині цілих чисел: $\exists k\ (y = kx)$. Часто такі означення записують у вигляді $x\,|\,y \equiv \exists k\ (y = kx)$. Замість знака рівносильності $\equiv$ пишуть також знак $\overset{\text{озн}}{\Leftrightarrow}$, який читається "за означенням".

За допомогою предиката $x\,|\,y$ можна природно означити унарний предикат "x — просте число" (позначимо його через $P(x)$):

$$P(x)\overset{\text{озн}}{\Leftrightarrow}\forall y\ ((y\,|\,x)\rightarrow((y=1)\vee(y=-1)\vee$$
$$\vee\ (y=x)\vee(y=-x)))\wedge(x>1).$$

Наведемо ще кілька прикладів означень із математичного аналізу. Відоме означення границі числової послідовності можна записати так:

$$\lim a_i = a \overset{\text{озн}}{\Leftrightarrow}\forall(\varepsilon>0)\ \exists(k\in N)\ \forall(i\in N\wedge i>k)\ (|a_i - a|<\varepsilon).$$

Аналогічно можна записати класичні означення різних варіантів поняття неперервності дійсної функції f:

1) функція $f(x)$ *неперервна в точці* a $\overset{\text{озн}}{\Leftrightarrow}$

$$\forall(\varepsilon>0)\ \exists(\eta>0)\ \forall(x\in R)\ ((|x-a|<\eta)\rightarrow(|f(a)-f(x)|<\varepsilon));$$

2) функція $f(x)$ *неперервна на інтервалі* $(a,b)\overset{\text{озн}}{\Leftrightarrow}$

$$\forall(c\in(a,b))\ \ \forall(\varepsilon>0)\ \exists(\eta>0)\ \forall(x\in(a,b))\ ((|x-c|<\eta)\rightarrow(|f(c)-f(x)|<\varepsilon));$$

3) функція $f(x)$ *рівномірно неперервна на інтервалі* $(a,\ b)$ $\overset{\text{озн}}{\Leftrightarrow}$

$$\forall(\varepsilon>0)\ \ \exists(\eta>0)\ \forall(c\in(a,b))\ \forall(x\in(a,b))\ ((|x-c|<\eta)\rightarrow(|f(c)-f(x)|<\varepsilon)).$$

Означення основних теоретико-множинних операцій і відношення включення для множин можна записати так:

$x\in A\cup B\overset{\text{озн}}{\Leftrightarrow}(x\in A)\vee(x\in B)$, $\quad x\in A\cap B\overset{\text{озн}}{\Leftrightarrow}(x\in A)\wedge(x\in B)$,

$x\in A\setminus B\overset{\text{озн}}{\Leftrightarrow}(x\in A)\wedge(x\notin B)$, $\quad A\subseteq B\overset{\text{озн}}{\Leftrightarrow}\forall x\ ((x\in A)\rightarrow(x\in B))$

тощо.

3.7. Задачі і вправи

1. Записати символікою логіки предикатів твердження (якщо для певного предиката немає загальновживаного позначення, то позначити його якоюсь предикатною буквою):

(а) кожне число, кратне 21, кратне 3 і 7;

(б) усі числа, кратні 60, кратні 10 і кратні 6, але не будь-яке число, кратне 10 і кратне 6, кратне 60;

(в) усі трансцендентні числа ірраціональні, але не всі ірраціональні числа трансцендентні.

2. Використовуючи символічну мову логіки предикатів, записати для множини N усіх натуральних чисел твердження:

(а) про замкненість операції додавання і операції множення;

(б) про незамкненість операції віднімання і операції ділення.

3. Застосовуючи символіку логіки предикатів, записати означення

(а) симетричної різниці множин A, B;

(б) декартового добутку множин A, B.

4. Застосовуючи символіку логіки предикатів, записати факт існування множини всіх підмножин (булеана) для кожної множини.

5. Нехай A і B — непорожні множини. Використовуючи символіку логіки предикатів, записати означення:

(а) відповідності;

(б) всюди визначеної відповідності;

(в) функціональної відповідності;

(г) сюр'єктивної відповідності;

(д) ін'єктивної відповідності;

(е) бієктивної відповідності

між A і B.

6. Нехай M — непорожня множина. Записати логіко-математичною символікою такі означення:

(а) відношення;

(б) рефлексивного відношення;

(в) антирефлексивного відношення;

(г) симетричного відношення;

(д) антисиметричного відношення;

(е) транзитивного відношення

на множині M.

7. Нехай M — частково впорядкована множина. Записати логіко-математичною символікою означення:

(а) верхньої грані;

(б) нижньої грані;

(в) точної верхньої грані;

(г) точної нижньої грані;

(д) найбільшого елемента;

(е) найменшого елемента;

(є) максимального елемента;

(ж) мінімального елемента

непорожньої підмножини A множини M.

8. Виходячи із законів елементарної математики, визначити, чи є запропоноване твердження істинним (предметна область — множина всіх дійсних чисел):

(а) $\forall x\ \forall y\ ((x = 0 \vee y = 1) \rightarrow y^x = 1)$;

(б) $\forall p\ \exists q\ \forall x\ (x^2 + px + q > 0)$;

(в) $\forall x\ \forall y\ (y^x = 1 \rightarrow (x = 0 \vee y = 1))$;

(г) $\exists q\ \forall p\ \forall x\ (x^2 + px + q > 0)$.

9. Оцінити істинність математичного твердження

$$\forall x\ \forall y\ \exists z\ (\neg(x = y) \rightarrow (x < z \wedge z < y \vee y < z \wedge z < x))$$

у випадку, коли універсальною множиною є множина:

(а) N натуральних чисел; (в) Q раціональних чисел;

(б) Z цілих чисел; (г) R дійсних чисел.

Яку властивість предметної області виражає це твердження?

10. Записати символічною мовою логіки предикатів із використанням символів відповідних предикатів доведення таких теоретико-множинних рівностей:

(а) $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C)$;

(б) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.

11. Із хибного твердження $\forall a\ \forall b\ \forall x\ (ax = bx \rightarrow a = b)$ вивести, що $2 \times 2 = 5$.

РОЗДІЛ 4

ТЕОРІЯ ГРАФІВ

Роком виникнення теорії графів одностайно вважають 1736, коли Леонард Ейлер опублікував розв'язання так званої задачі про кенігсберзькі мости, а також знайшов загальний критерій існування ейлерового циклу в графі (див. розд. 4.12).

Отримання дальших істотних результатів у цій галузі датують серединою XIX століття. Однак початок проведення активних систематичних досліджень і становлення теорії графів як окремого авторитетного розділу сучасної математики відбулося ще майже 100 років по тому, тобто в середині XX століття. Саме з цього часу граф став однією з найпоширеніших і найпопулярніших математичних моделей у багатьох сферах науки та техніки. Картинка у вигляді набору точок на площині та ліній, проведених між деякими з них, стала зручною і наочною формою зображення найрізноманітніших об'єктів, процесів і явищ (див. розд. 4.14).

Великою мірою це пов'язано з виникненням, бурхливим розвитком і поширенням електронних обчислювальних машин і, як наслідок, значним зростанням ролі задач дискретного характеру. Математика від "обслуговування" переважно фізики переходить до впровадження своїх методів у інші сфери людської діяльності. Одним із потужних інструментів такого проникнення є граф.

Із суто формальної точки зору граф можна розглядати як один із різновидів алгебричної системи (а саме, як модель), а отже, і всю теорію графів — як розділ сучасної алгебри. Справді, результати й методи алгебри широко використовують у теорії графів. Однак за останні півстоліття активного інтенсивного й екстенсивного розвитку теорія графів виробила свою достатньо специфічну власну проблематику і методологію. На сьогодні теорія графів — одна зі складових математичного апарату кібернетики, важливий розділ дискретної математики. Тому вона входить до навчальних програм університетів та інших закладів освіти як окрема дисципліна, або як розділ курсу "Дискретна математика".

Безумовно, в одному невеликому розділі неможливо викласти весь спектр досягнень сучасної теорії графів. Нашою метою є ознайомлення з основами теорії графів, деякими класичними проблемами теорії та методами їх розв'язання. Цей вступний курс має полегшити зацікавленому читачеві подальше знайомство з результатами та різноманітними застосуваннями теорії графів за допомогою спеціальних монографій і підручників [12; 16; 21; 23–25; 29; 31; 33 та ін.].

Крім традиційних (можна сказати, класичних) розділів теорії графів до матеріалу глави включено фрагмент теми (див. розд. 4.7), яку відносять до прикладної теорії алгоритмів і називають "Алгоритми на графах", або ширше — "Комбінаторні алгоритми" [16; 21; 24].

Дещо спрощено відмінність різних підходів до однієї й тієї самої проблеми, пов'язаної з графами, можна продемонструвати на прикладі задачі перевірки ізоморфності двох заданих графів. Класична (абстрактна) теорія алгоритмів констатує, що ця проблема розв'язна. Класична теорія графів уточнює, що для її розв'язання потрібно перебрати й перевірити не більше ніж $n!$ варіантів (n — кількість вершин в обох графах). Прикладна теорія алгоритмів (і зокрема, її розділ під назвою "Алгоритми на графах") займається конструюванням нетривіальних процедур, що дають змогу визначити, чи ізоморфні задані графи, і в більшості випадків здійснити цю перевірку достатньо ефективно, істотно скоротивши перебір можливих варіантів.

Розділ "Алгоритми на графах" є своєрідним містком між класичною теорією графів та прикладною теорією алгоритмів і програмуванням. Цей місток із кожним роком розширюється й міцніє. Деякі автори вже сьогодні відносять розділ "Алгоритми на графах" до загальної теорії графів [16; 21].

4.1. Поняття графа. Способи задання графів

Нехай V — непорожня скінченна множина, а $V^{(2)}$ — множина всіх двохелементних підмножин (невпорядкованих пар різних елементів) множини V.

Графом (неорієнтованим графом) G називається пара множин (V, E), де E — довільна підмножина множини $V^{(2)}$ ($E \subseteq V^{(2)}$); позначається $G = (V, E)$.

Елементи множини V називають ***вершинами*** графа G, а елементи множини E — його ***ребрами***. Відповідно V називається ***множиною вершин*** і E — ***множиною ребер*** графа G.

Традиційно ребро $\{v, w\}$ записують за допомогою круглих дужок (v, w) (іноді просто vw).

Граф, який складається з однієї вершини, називається ***тривіальним***.

Оскільки для тривіального графа чи так званих ***порожніх*** графів $G = (V, \varnothing)$ неважко перевірити переважну більшість властивостей і тверджень, то надалі не будемо кожен раз, формулюючи та доводячи ті чи інші загальні твердження теорії графів, спеціально обумовлювати, що йдеться про нетривіальні графи (при цьому для тривіального або порожнього графів результат може бути дещо іншим).

Нехай задано граф $G = (V, E)$. Якщо $(v, w) \in E$, то кажуть, що ***вершини*** v і w ***суміжні***, інакше вершини v і w ***несуміжні***. Якщо $e = (v, w)$ — ребро графа, то вершини v і w називають ***кінцями*** ребра e. У цьому разі кажуть також, що ребро e ***з'єднує*** вершини v і w. Вершину v і ребро e називають ***інцидентними***, якщо v — кінець ребра e. Два ***ребра*** називають ***суміжними***, якщо вони мають спільну вершину.

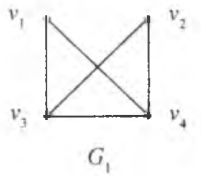
Існує кілька ***способів задання*** графів.

Один із них — це задати кожну з множин V та E за допомогою переліку їх елементів.

Приклад 4.1. $G_1 = (V_1, E_1)$, де $V_1 = \{v_1, v_2, v_3, v_4\}$ й $E_1 = \{(v_1, v_3), (v_1, v_4), (v_2, v_3), (v_2, v_4), (v_3, v_4)\}$, — це граф із чотирма вершинами та п'ятьма ребрами, а $G_2 = (V_2, E_2)$, де $V_2 = \{v_1, v_2, v_3, v_4, v_5\}$ і $E_2 = \{(v_1, v_2), (v_2, v_4), (v_1, v_5), (v_3, v_2), (v_3, v_5), (v_4, v_1), (v_5, v_4)\}$ — граф із п'ятьма вершинами та сімома ребрами.

Граф $G = (V, E)$ зручно зображати за допомогою рисунка на площині, який називають ***діаграмою*** графа G. Вершинам графа G поставлено в бієктивну відповідність точки площини; точки, що відповідають вершинам v і w, з'єднують лінією (відрізком або кривою) тоді й тільки тоді, коли v і w — суміжні вершини. Зрозуміло, що діаграма графа змінює свій вигляд залежно від вибору відповідних точок на площині.

Приклад 4.2. На рисунку зображено діаграми графів G_1 і G_2 з попереднього прикладу (рис. 4.1).



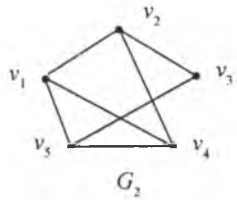


Рис. 4.1

Графи можна задавати також за допомогою матриць.

Занумеруємо всі вершини графа G натуральними числами від 1 до n. ***Матрицею суміжності*** A графа G називається квадратна $n \times n$-матриця, у якій елемент a_{ij} i-го рядка та j-го стовпчика дорівнює 1, якщо вершини v_i та v_j з номерами i та j суміжні, і дорівнює 0 в іншому разі.

Приклад 4.3. Для графів G_1 і G_2 маємо відповідно

$$A_1 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Очевидно, що матриці суміжності графів симетричні.

Занумеруємо всі вершини графа G числами від 1 до n, а всі його ребра — числами від 1 до m. ***Матрицею інциденності*** B графа G називається $n \times m$-матриця, у якій елемент b_{ij} i-го рядка та j-го стовпчика дорівнює 1, якщо вершина v_i з номером i інцидентна ребру e_j з номером j, і дорівнює 0 в іншому разі.

Приклад 4.4. Для графів G_1 і G_2 маємо такі матриці інцидентності (ребра графів занумеровано в тому порядку, у якому їх виписано у прикладі 4.1):

$$B_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \quad B_2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Нарешті, ще одним способом задання графів є ***списки суміжності***. Кожній вершині графа відповідає свій список. У список, що відповідає вершині v, послідовно записують усі суміжні їй вершини.

Приклад 4.5. Для графів G_1 і G_2 маємо такі списки суміжності:

G_1:	G_2:
v_1: v_3, v_4	v_1: v_2, v_4, v_5
v_2: v_3, v_4	v_2: v_1, v_3, v_4
v_3: v_1, v_2, v_4	v_3: v_2, v_5
v_4: v_1, v_2, v_3	v_4: v_1, v_2, v_5
	v_5: v_1, v_3, v_4.

Вибір і зручність того чи іншого способу задання графів залежить від особливостей розв'язуваної задачі.

4.1. Задачі і вправи

1. Нехай задано граф $G = (V, E)$:

(а) $V = \{1, 2, 3, 4\}$, $E = \{(1, 3), (2, 3), (3, 4), (4, 1), (4, 2)\}$;

(б) $V = \{a, b, c, d, e\}$, $E = \{(a, d), (b, c), (b, e), (c, e), (d, b), (d, e), (e, a)\}$;

(в) $V = \{1, 2, 3, 4\}$, $E = \{(1, 2), (2, 3), (3, 4)\}$;

(г) $V = \{A, B, C, D\}$, $E = \{(A, B), (A, D), (B, C), (B, D), (C, A), (C, D)\}$.

Побудувати діаграму, матриці суміжності й інцидентності для кожного із цих графів.

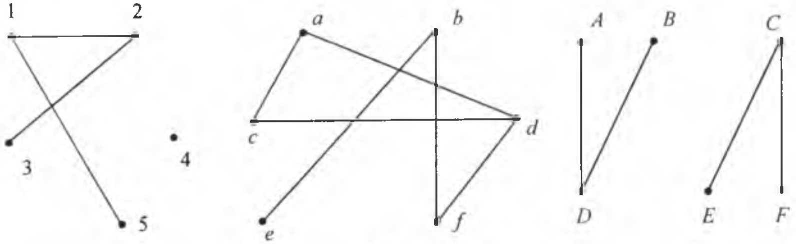
2. Нехай $V = \{a, b, c, d, e\}$. Граф $G = (V, E)$ задано за допомогою матриці суміжності A:

$$\text{(а) } A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}; \quad (\ )\ A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix};$$

$$(\ )\ A = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}; \quad (\ )\ A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Визначити множину ребер E графа G. Побудувати діаграму і матрицю інцидентності.

3. Граф G задано його діаграмою



а *б* *в*

Визначити множину вершин V і множину ребер E, матриці суміжності й інцидентності графа G.

4. Нехай задано матрицю суміжності A деякого графа G. Як за допомогою матриці A визначити:

(а) кількість вершин графа G;

(б) кількість ребер графа G;

(в) матрицю інцидентності графа G?

5. Нехай A — матриця суміжності, а B — матриця інцидентності графа G. Довести, що матриця A дорівнює матриці BB^{T}, в якій усі діагональні елементи замінено нулями.

6. Довести, що в довільному графі G із шістьма вершинами завжди знайдуться три вершини, які або попарно суміжні, або попарно несуміжні.

4.2. Підграфи. Ізоморфізм графів. Алгебра графів

Граф $G_1 = (V_1, E_1)$ називається ***підграфом*** графа $G = (V, E)$, якщо $V_1 \subseteq V$ і $E_1 \subseteq E$.

Важливі класи становлять підграфи, які можна отримати в результаті застосування до заданого графа операцій вилучення вершини або вилучення ребра.

Операція вилучення вершини v з графа $G = (V, E)$ полягає у вилученні з множини V елемента v, а з множини E — усіх ребер, інцидентних v.

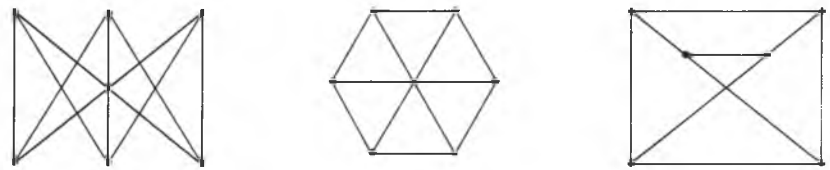
Операція вилучення ребра e з графа $G = (V, E)$ — це вилучення елемента e з множини E. При цьому всі вершини зберігаються.

Графи $G_1 = (V_1, E_1)$ і $G_2 = (V_2, E_2)$ називаються ***ізоморфними***, якщо існує таке взаємно однозначне відображення φ множини вершин V_1 на множину вершин V_2, що ребро $(v, w) \in E_1$ тоді й тільки тоді, коли ребро $(\varphi(v), \varphi(w)) \in E_2$. Відображення φ називається ***ізоморфним відображенням***, або ***ізоморфізмом***, графа G_1 на граф G_2.

Таким чином, ізоморфні графи відрізняються фактично лише ідентифікаторами (іменами) своїх вершин. З погляду теорії графів ця відмінність неістотна, тому зазвичай ізоморфні графи ототожнюють і, зображаючи графи у вигляді діаграм, або зовсім не ідентифікують їх вершини, або нумерують вершини натуральними числами.

Ізоморфне відображення графа G на себе називається ***автоморфізмом*** графа G. Автоморфізм φ графа $G = (V, E)$, при якому для кожної вершини $v \in V$ виконується рівність $\varphi(v) = v$ (тобто φ — тотожне відображення), називається ***тривіальним автоморфізмом***.

Приклад 4.6. Пропонуємо переконатися, що всі графи, зображені на рис. 4.2, ізоморфні між собою, а графи на рис. 4.3 — не ізоморфні.



G_1 G_2 G_3

Рис. 4.2

H_1 H_2

Рис. 4.3

Відношення ізоморфізму є відношенням еквівалентності на сукупності графів.

Теорема 4.1. Графи G_1 і G_2 ізоморфні тоді й тільки тоді, коли матрицю суміжності (інцидентності) одного з цих графів можна одержати з матриці суміжності (інцидентності) іншого графа за допомогою відповідних перестановок рядків і стовпчиків.

Доведення. Справді, як було зазначено вище, ізоморфні графи G_1 і G_2 відрізняються між собою лише порядком нумерації вершин, тобто існує бієктивне відображення φ множини номерів вершин першого графа на множину номерів вершин другого. Отже, кожен елемент $a_{ij}^{(1)}$ матриці суміжності A_1 графа G_1 збігається з елементом $a_{\varphi(i)\varphi(j)}^{(2)}$ (тобто елементом, який стоїть у рядку з номером $\varphi(i)$ і стовпчику з номером $\varphi(j)$) матриці суміжності A_2 графа G_2. Послідовно одночасно міняючи місцями (переставляючи) рядки і стовпчики з номерами i та $\varphi(i)$ для всіх $i = 1, 2, \ldots, n$, матрицю суміжності A_1 можна перетворити в матрицю суміжності A_2, і навпаки.

Якщо відображення φ відоме, то таке перетворення виконати неважко. Коли ж за допомогою матриць суміжності потрібно перевірити, чи ізоморфні два задані графи з n вершинами кожний, треба здійснити різноманітні одночасні перестановки рядків і стовпчиків однієї з них. Якщо після чергової перестановки дістанемо матрицю, яка повністю збігається з другою, то ці графи ізоморфні. Однак щоб у такий спосіб з'ясувати, що задані графи неізоморфні, потрібно виконати всі $n!$ перестановок рядків і стовпчиків. Уже для порівняно невеликих значень n здійснити такий перебір практично неможливо навіть за допомогою обчислювальної машини. У прикладній теорії алгоритмів розроблено різноманітні алгоритми перевірки ізоморфізму графів, які для більшості графів (або окремих їх типів) дають змогу істотно скоротити обсяг потрібних перевірок [16; 21; 24].

Для матриць інцидентності графів G_1 і G_2 з n вершинами і m ребрами кожний справедливі аналогічні міркування. Відмінність полягає в тому, що коли графи G_1 і G_2 ізоморфні, то для множин їх вершин існує бієкція φ, а для множин ребер — інша бієкція ψ. Загальна ж кількість потрібних кроків для перевірки ізоморфізму графів G_1 і G_2 в цьому разі не перевищує $n!m!$.

Граф $G = (V, E)$ називається ***повним***, якщо будь-які дві його вершини суміжні (тобто $E = V^{(2)}$). Повний граф з n вершинами позначають K_n.

Будь-яка підстановка множини вершин повного графа K_n є автоморфізмом цього графа. Тому кількість усіх можливих автоморфізмів графа K_n дорівнює $n!$.

Для графів можна означити операції об'єднання, перетину та доповнення.

Об'єднанням графів $G_1 = (V_1, E_1)$ і $G_2 = (V_2, E_2)$ називається граф $G = (V_1 \cup V_2, E_1 \cup E_2)$; позначається $G = G_1 \cup G_2$.

Об'єднання $G = G_1 \cup G_2$ називають ***прямою сумою*** графів G_1 і G_2, якщо $V_1 \cap V_2 = \varnothing$.

Перетином і ***різницею*** графів $G_1 = (V, E_1)$ і $G_2 = (V, E_2)$ з однаковими множинами вершин називають відповідно графи $G' = (V, E_1 \cap E_2)$ і $G'' = (V, E_1 \setminus E_2)$; позначають $G' = G_1 \cap G_2$ та $G'' = G_1 \setminus G_2$.

Доповненням графа $G = (V, E)$ називається граф $\overline{G} = (V, V^{(2)} \setminus E)$. Отже, граф $\overline{G}$ має ту саму множину вершин V, що й граф G, а вершини графа $\overline{G}$ суміжні тоді й лише тоді, коли вони несуміжні в G. Для графа G з n вершинами $\overline{G} = K_n \setminus G$.

Так можна означити алгебру графів $A = \langle \Gamma, \{\cup, \cap, {}^{-}\} \rangle$ (типу (2, 2, 1)), носієм якої є множина Γ всіх графів. Існують й інші операції для графів, отже, сигнатуру алгебри A можна розширювати.

Неважко переконатись, що справджується таке твердження.

Теорема 4.2. Графи G_1 і G_2 ізоморфні тоді й тільки тоді, коли ізоморфні їх доповнення $\overline{G}_1$ і $\overline{G}_2$.

Приклад 4.7. Об'єднання та перетин графів H_1 і H_2 з попереднього прикладу зображено на рис. 4.4, а доповнення графів G_2 та H_2 — на рис. 4.5.

$H_1 \cup H_2$ $H_1 \cap H_2$

Рис. 4.4

$\overline{G}_2$ $\overline{H}_2$

Рис. 4.5

Граф G, ізоморфний своєму доповненню $\overline{G}$, називається ***самодоповнювальним***.

Наприклад, граф $G = (V, E)$, в якому $V = \{a, b, c, d\}$ і $E = \{(a, b), (b, c), (c, d)\}$, самодоповнювальний.

4.2. Задачі і вправи

1. Накресліть діаграму повного графа з n вершинами K_n для
(а) $n = 2$; (б) $n = 3$; (в) $n = 4$; (г) $n = 5$.

2. Скільки ребер містить повний граф з n вершинами?

3. Чи існує повний граф, у якого кількість ребер дорівнює:
(а) 15; (в) 199…900…0 (n дев'яток і n нулів);
(б) 18; (г) $8k^2 + 2k$, $k \in N$?

4. Чому дорівнює кількість ребер у графі $\overline{G}$, якщо граф G має n вершин і k ребер?

5. Довести, що для довільного графа G об'єднання $G \cup \overline{G}$ — повний граф.

6. Нехай графи $G_1 = (V, E_1)$ і $G_2 = (V, E_2)$ задано за допомогою матриць суміжності A_1 і A_2 відповідно ($|V| = n$). Визначити матрицю суміжності A для графа:
(а) $G_1 \cup G_2$; (б) $G_1 \cap G_2$; (в) $G_1 \setminus G_2$; (г) $\overline{G_2}$.

7. Довести, що ізоморфне відображення φ графа $G_1 = (V_1, E_1)$ на граф $G_2 = (V_2, E_2)$ встановлює певне взаємно однозначне відображення ψ множини ребер E_1 на множину ребер E_2.

8. Довести, що ізоморфні графи мають однакову кількість вершин і однакову кількість ребер.

9. Довести, що ізоморфізм є відношенням еквівалентності на сукупності графів.

10. Довести, що графи ізоморфні тоді й тільки тоді, коли ізоморфні їх доповнення.

11. Знайти нетривіальний самодоповнювальний граф із найменшою кількістю вершин.

12. Довести, що кількість вершин будь-якого самодоповнювального графа дорівнює $4k$ чи $4k + 1$, $k \in N$.

13. Довести, що довільний самодоповнювальний граф містить $4k^2 - k$ чи $4k^2 + k$ ребер, $k \in N$.

14. Побудувати нетривіальний граф із найменшою кількістю вершин, який не має нетривіальних автоморфізмів.

4.3. Графи та бінарні відношення

Між множиною всіх графів із множиною вершин V та множиною всіх антирефлексивних симетричних бінарних відношень на V існує взаємно однозначна відповідність: графу $G = (V, E)$ відповідає таке відношення R на V, що $(v, w) \in R$ тоді й тільки тоді, коли $(v, w) \in E$, $v, w \in V$. Зокрема, порожньому графу $G = (V, \varnothing)$ відповідає порожнє відношення на V ($R = \varnothing$), а повному графу — відношення $(V \times V) \setminus i_V$, де i_V — діагональне (тотожне) відношення: $i_V = \{(v, v) \mid v \in V\}$.

Операціям об'єднання, перетину та доповнення графів відповідають аналогічні операції для відношень. Якщо графам G_1 і G_2 відповідають відношення R_1 і R_2, то графам $G_1 \cup G_2$, $G_1 \cap G_2$ та $\overline{G_1}$ відповідають відношення $R_1 \cup R_2$, $R_1 \cap R_2$ та $\overline{R_1} \setminus i_V = (V \times V) \setminus (R_1 \cup i_V)$.

Якщо R — транзитивне відношення, то у відповідному графі G для кожної пари ребер (v, w), $(w, u) \in E$ існує замикальне ребро $(v, u) \in E$.

4.4. Степені вершин графа

Степенем $\delta(v)$ ***вершини*** v називається кількість інцидентних їй ребер. Вершина степеня 0 називається ***ізольованою***, а вершина степеня 1 — ***кінцевою*** (або ***висячою***).

Кубічним графом називається граф, степені всіх вершин якого дорівнюють 3.

Нижченаведені твердження містять прості, але важливі властивості графів.

Теорема 4.3. У будь-якому графі $G = (V, E)$

$$\sum_{v \in V} \delta(v) = 2|E|. \tag{4.1}$$

Справедливість цього твердження випливає з того, що кожне ребро графа інцидентне двом вершинам, тому воно вносить у суму степенів усіх вершин рівно дві одиниці.

Теорема 4.4. У будь-якому графі $G = (V, E)$ кількість вершин непарного степеня парна.

Доведення. Нехай $V_п$ — множина вершин парного степеня, а $V_н$ — множина вершин непарного степеня графа G. Зрозуміло, що $V_п \cup V_н = V$ та $V_п \cap V_н = \varnothing$, тоді

$$\sum_{v \in V} \delta(v) = \sum_{v \in V_п} \delta(v) + \sum_{v \in V_н} \delta(v) = 2k + \sum_{v \in V_н} \delta(v) = 2|E|.$$

Отже, сума $\sum_{v\in V}\delta(v) = 2(|E| - k)$ — парне число. Оскільки кожен із доданків $\delta(v)$ цієї суми — непарне число, то кількість цих доданків парна.

Відомі такі інтерпретації наведених теорем. Загальна кількість рукостискань, які роблять люди у будь-якій спільноті, завжди парна. Крім того, кількість людей цієї спільноти, що зробили непарну кількість рукостискань, завжди є також парною.

Користуючись теоремою 4.3, неважко довести таке твердження.

Лема 4.1. Кількість ребер у повному графі K_n з n вершинами дорівнює $n(n-1)/2$.

Кожна з n вершин повного графа K_n має степінь $n - 1$. Отже, $\sum_{v\in V}\delta(v) = n(n-1) = 2|E|$.

4.3. Задачі і вправи

1. Чому дорівнює степінь кожної вершини в повному графі з n вершинами?

2. Як визначити степінь певної вершини графа G за його матрицею суміжності A чи матрицею інцидентності B?

3. Нехай A — матриця суміжності графа G з n вершинами. Довести, що i-й діагональний елемент матриці A^2 дорівнює степеню $\delta(v_i)$ i-ї вершини графа G, $i = 1, 2, \ldots, n$.

4. Нехай B — матриця інцидентності графа G з n вершинами. Довести, що i-й діагональний елемент матриці BB^{T} дорівнює степеню $\delta(v_i)$ i-ї вершини графа G, $i = 1, 2, \ldots, n$.

5. Скільки вершин має бути в кубічному графі? Скільки ребер у такому графі?

6. Довести, що доповнення жодного кубічного графа не є кубічним графом.

7. Довести, що в будь-якому графі з n вершинами ($n \ge 2$) завжди знайдуться принаймні дві вершини з однаковими степенями.

8. У графі з n вершинами тільки дві вершини мають однакові степені. Чи можуть обидві ці вершини мати степінь 0 або $n - 1$?

9. У футбольному турнірі беруть участь 17 команд. Довести, що в будь-який момент є команда, яка зіграла парну кількість матчів.

10. Довести, що в ізоморфних графів кількість вершин степеня k однакова для довільного k.

4.5. Шлях у графі. Зв'язність графів

Маршрутом (або ***шляхом***) у графі $G = (V, E)$ називається послідовність

$$v_1, e_1, v_2, e_2, \ldots, e_k, v_{k+1} \qquad (4.2)$$

вершин v_i і ребер e_i така, що кожні два сусідні ребра в цій послідовності суміжні, отже $e_i = (v_i, v_{i+1})$, $i = 1, 2, \ldots, k$. Вершина v_1 називається ***початком***, а вершина v_{k+1} — ***кінцем*** шляху. Усі інші вершини цього шляху називаються ***проміжними***, або ***внутрішніми***.

Кількість k ребер у маршруті називається його ***довжиною***. Кажуть, що цей маршрут ***з'єднує*** вершини v_1 і v_{k+1}, або ***веде*** з вершини v_1 у вершину v_{k+1}.

Маршрутом довжини 0 вважають послідовність, що складається з єдиної вершини.

Маршрут, у якому всі ребра попарно різні, називається ***ланцюгом***, а той, у якому всі вершини попарно різні, — ***простим ланцюгом***.

Маршрут (4.2) називається ***замкненим*** (або ***циклічним***), якщо $v_1 = v_{k+1}$. Замкнений ланцюг називається ***циклом***, а замкнений простий ланцюг — ***простим циклом***.

Лема 4.2. Будь-який маршрут, що веде з вершини v у вершину w ($v \ne w$), містить у собі простий ланцюг, що веде з v в w.

Доведення. Справді, нехай $v, e_1, v_2, e_2, \ldots, e_k, w$ — маршрут M, що веде з v в w і такий, що серед його проміжних вершин є однакові. Якщо $v_i = v_j$ ($i < j$), то, викинувши з M ділянку (циклічний маршрут) від v_i до v_j, отримаємо маршрут M' — $v, e_1, v_2, e_2, \ldots, v_i, e_{j+1}, \ldots, e_k, w$, який також веде з v у w. Якщо M' — не простий ланцюг, то процедуру вилучення його внутрішніх циклічних ділянок можна повторити. Врешті-решт отримаємо простий ланцюг, що з'єднує вершини v і w.

Граф, усі ребра якого утворюють простий цикл довжини n, позначають C_n. Простий цикл довжини 3 називається ***трикутником***.

Граф називається ***зв'язним***, якщо будь-яку пару його вершин можна з'єднати деяким маршрутом.

Компонентою зв'язності (або ***зв'язною компонентою***) графа G називається його зв'язний підграф, що не є власним підграфом жодного іншого зв'язного підграфа графа G.

Відстанню між вершинами v та w зв'язного графа (позначається $d(v, w)$) називається довжина найкоротшого простого ланцюга, що з'єднує вершини v і w.

Оскільки кожна вершина графа $G = (V, E)$ з'єднана сама з собою маршрутом довжини 0, то для всіх $v \in V$ виконується рівність $d(v, v) = 0$.

Означена функція відстані задовольняє три аксіоми метрики, тобто для будь-яких вершин $v, w, u \in V$ виконується

1) $d(v, w) \geq 0$; $d(v, w) = 0$ тоді й тільки тоді, коли $v = w$;
2) $d(v, w) = d(w, v)$;
3) $d(v, w) \leq d(v, u) + d(u, w)$.

Справедливість перших двох аксіом очевидна. Доведемо третю аксіому, яка називається *нерівністю трикутника*. Нехай $v, e_1, v_2, e_2, \ldots, e_k, u$ — маршрут із v в u, а $u, e_1', u_2, e_2', \ldots, e_l', w$ — маршрут з u у w. Тоді послідовність $v, e_1, v_2, e_2, \ldots, e_k, u, e_1', u_2, e_2', \ldots, e_l', w$ — це маршрут, що веде з v у w і має довжину $d(v, u) + d(u, w)$. Зрозуміло, що цей маршрут не може бути коротшим від найкоротшого маршруту, який веде з v у w, тому $d(v, w) \leq d(v, u) + d(u, w)$.

Ексцентриситетом $e(v)$ вершини v зв'язного графа $G = (V, E)$ називається найбільша з відстаней між вершиною v і всіма іншими вершинами графа G, тобто $e(v) = \max\limits_{w \in V}\{d(v, w)\}$.

Діаметром зв'язного графа G (позначається $D(G)$) називається максимальний з усіх ексцентриситетів вершин графа G. Мінімальний з усіх ексцентриситетів вершин зв'язного графа G називають його ***радіусом*** і позначають $R(G)$.

Вершину v називають ***центральною***, якщо $e(v) = R(G)$. ***Центром*** графа G називається множина всіх його центральних вершин.

Вершини v і w графа G називаються ***зв'язаними***, якщо в G існує маршрут, що їх з'єднує.

Відношення зв'язаності Z рефлексивне, транзитивне та симетричне, тобто є відношенням еквівалентності на множині V. Розглянемо фактор-множину $V/Z = \{V_1, V_2, \ldots, V_t\}$. Підграфи $G_i = (V_i, E_i)$, де $E_i = E \cap V_i^{(2)}$, є компонентами зв'язності графа G. Крім того, виконується $G = G_1 \cup G_2 \cup \ldots \cup G_t$. Цей факт можна сформулювати у вигляді такого твердження.

Теорема 4.5. Будь-який граф можна однозначно зобразити у вигляді прямої суми його компонент зв'язності.

Якщо граф G зв'язний, то всі його вершини попарно зв'язані, тобто $V/Z = \{V\}$ і G має єдину зв'язну компоненту, яка збігається із самим графом G.

Теорема 4.6. Для будь-якого графа $G = (V, E)$ він сам або його доповнення $\bar{G}$ є зв'язним графом.

Доведення. Якщо G — зв'язний граф, то твердження теореми виконано.

Нехай $G = (V, E)$ — незв'язний граф і $G_1 = (V_1, E_1)$ — одна з його компонент зв'язності. Розглянемо графи G_1 і $G' = (V', E')$, де $V' = V \setminus V_1$ і $E' = E \setminus E_1$. Для будь-якої пари вершин $v \in V_1$ і $w \in V'$ у графі $\bar{G}$ існує ребро (v, w), тому що ці вершини несуміжні у графі G. Оскільки для кожної пари вершин $v_1, v_2 \in V_1$ графа G_1 і довільної вершини $w \in V'$ існують ребра (v_1, w) і (v_2, w), які належать множині ребер графа $\bar{G}$, то у графі $\bar{G}$ такі вершини v_1 і v_2 зв'язані. Аналогічно встановлюємо зв'язність у графі $\bar{G}$ будь-якої пари вершин w_1 і w_2 з множини V'. Отже, усі пари вершин графа $\bar{G}$ зв'язані між собою.

Наслідок 4.6.1. Якщо G — незв'язний граф, то граф $\bar{G}$ зв'язний і $D(\bar{G}) \leq 2$.

Справді, якщо G — незв'язний граф, то з доведення теореми випливає, що $\bar{G}$ зв'язний, і для будь-яких двох вершин v та w графа $\bar{G}$ виконується або $d(v, w) = 1$, або $d(v, w) = 2$.

Доведена теорема дає змогу звести розв'язання деяких важливих проблем теорії графів до випадку зв'язних графів. Зокрема, це стосується проблеми встановлення ізоморфізму заданих графів G_1 і G_2 (див. теорему 4.2).

Теорема 4.7. Нехай $G = (V, E)$ — зв'язний граф, а e — деяке його ребро. Розглянемо граф G', отриманий з G вилученням ребра e.

а). Якщо ребро e належить деякому циклу графа G, то граф G' зв'язний.

б). Якщо ребро e не належить жодному циклу графа G, то граф G' незв'язний і має рівно дві компоненти зв'язності.

Доведення. а). Розглянемо довільні дві вершини v і w графа G'.

Якщо маршрут M, що з'єднує вершини v та w у зв'язному графі G, не містить ребра e, то він з'єднуватиме вершини v і w й у графі G'. Якщо ж ребро e належить маршруту M і $e = (u_1, u_2)$, то маршрут, що веде з v в w у графі G', можна побудувати так: беремо маршрут, що веде з v в u_1, додаємо до нього ту частину циклу, що містив ребро e, яка залишилась у графі G' і з'єднує вершини u_1 і u_2; відтак завершуємо його маршрутом з u_2 в w. Отже, граф G' зв'язний.

б). Нехай ребро $e = (u_1, u_2)$ не належить жодному циклу графа G. Тоді у графі G' вершини u_1 та u_2 будуть незв'язаними й належатимуть двом різним компонентам зв'язності G_1 і G_2 графа G'. Крім того, у графі G' стануть незв'язаними ті й тільки ті вершини, які були з'єднані у графі G маршрутом, що містив ребро e. Отже, кожна вершина v у G' зв'язана або з вершиною u_1, або з вершиною u_2, тобто v належатиме або G_1, або G_2.

Теорема 4.8. Нехай $G = (V, E)$ — граф з n вершинами та k компонентами зв'язності. Тоді число його ребер m задовольняє такі нерівності: $n - k \leq m \leq (n - k)(n - k + 1)/2$.

Доведення. Нижню оцінку $m \geq n - k$ доведемо індукцією за кількістю ребер у графі G.

Якщо $m = 0$, то $n = k$, і нерівність виконується.

Припустімо, що для всіх графів із числом ребер $m \leq t$ ($t \geq 0$) відповідна нерівність виконується. Розглянемо граф G з n вершинами і k компонентами зв'язності, який містить $t + 1$ ребро. Вилучимо з графа G довільне ребро. Тоді згідно з теоремою 4.7 отримаємо граф G' з n вершинами, t ребрами і кількістю компонент зв'язності, яка дорівнює k або $k + 1$. За припущенням індукції виконується або $t \geq n - k$, або $t \geq n - (k + 1)$. Для обох випадків $t + 1 \geq n - k$. Отже, бажана нерівність для графа G виконується.

Доведемо верхню оцінку. Розглянемо довільний граф G, що має n вершин, k компонент зв'язності та максимально можливу кількість ребер m. Тоді всі його зв'язні компоненти є повними графами. Нехай K_t й K_s — такі дві компоненти зв'язності графа G, що $t \geq s > 1$. Виконаємо таке перетворення графа G: перенесемо одну вершину v з компоненти K_s у компоненту K_t, тобто вилучимо вершину v з K_s і додамо до K_t, з'єднавши її з усіма вершинами K_t. Дістанемо граф G' з n вершинами, k компонентами зв'язності та кількістю ребер $m' = m - (s - 1) + t = m + (t - s + 1) > m$. Остання нерівність суперечить припущенню про максимальність числа ребер у графі G. Отже, шуканий граф з n вершинами, k компонентами зв'язності й найбільшою кількістю ребер має таку структуру: $k - 1$ його компоненти зв'язності є тривіальними графами, а остання компонента — повний граф з $n - k + 1$ вершиною. Кількість ребер у такому графі дорівнює $(n - k)(n - k + 1)/2$ (див. лему 4.1).

Наслідок 4.8.1. Довільний зв'язний граф з n вершинами містить не менше ніж $n - 1$ ребро.

Наслідок 4.8.2. Якщо у графі G з n вершинами кількість ребер більша ніж $(n - 1)(n - 2)/2$, то граф G зв'язний.

4.4. Задачі і вправи

1. Спростувати таке твердження: якщо деякий ланцюг, що веде з вершини v у вершину w, проходить через вершину u ($u \neq v$ та $u \neq w$), то він містить простий ланцюг, що веде з v у w і проходить через u.

2. Довести, що будь-який найкоротший ланцюг, що веде з вершини v у вершину w ($v \neq w$), є простим ланцюгом.

3. Довести, що у графі, степені всіх вершин якого більші 1, є цикл.

4. Довести, що у зв'язному графі будь-які два прості ланцюги максимальної довжини мають принаймні одну спільну вершину. Чи правильно, що вони завжди мають спільне ребро?

5. Довести, що коли у графі G тільки дві вершини v і w мають непарні степені, тоді ці вершини зв'язані у графі G.

6. Побудувати граф, центр якого:

(а) складається тільки з однієї вершини;

(б) складається тільки з двох вершин;

(в) складається з трьох вершин і не збігається з множиною всіх вершин;

(г) збігається з множиною всіх вершин.

7. Довести, що для довільного графа G виконуються нерівності $R(G) \leq D(G) \leq 2R(G)$.

8. Нехай для діаметра зв'язного графа G виконується нерівність $D(G) \geq 3$. Довести, що граф $\bar{G}$ є зв'язним і $D(\bar{G}) \leq 3$.

9. Довести, що самодоповнювальний граф завжди зв'язний.

10. Довести, що діаметр довільного самодоповнювального графа дорівнює або 2, або 3.

11. Довести, що в ізоморфних графів кількість простих циклів довжини l однакова для всіх l.

12. Знайти по три пари неізоморфних графів G_1 і G_2, у яких:

(а) кількість вершин степеня k однакова для всіх $k \geq 0$;

(б) кількість простих циклів довжини l однакова для всіх l;

(в) виконуються обидві умови з пунктів (а) і (б).

4.6. Перевірка зв'язності графів

Зв'язність заданого графа G зручно перевіряти за допомогою його матриці суміжності A.

Теорема 4.9. Нехай A — матриця суміжності графа $G = (V, E)$ з n вершинами ($|V| = n$). Тоді елемент $a_{ij}^{(k)}$ i-го рядка та j-го стовпчика матриці A^k дорівнює кількості шляхів довжини k, які ведуть у графі G з вершини v_i у вершину v_j.

Доведення проведемо індукцією за k.

Для $k = 1$ $a_{ij}^{(1)} = a_{ij}$. За означенням матриці A $a_{ij} = 1$ тоді й тільки тоді, коли у графі G є ребро, що веде з вершини v_i у вершину v_j. Але єдиний можливий шлях довжини 1 з v_i у v_j — це ребро (v_i, v_j). Отже, $a_{ij}^{(1)}$ дорівнює кількості шляхів довжини 1 з v_i у v_j.

Припустімо, що твердження теореми справджується для $k = m - 1$, $m \geq 2$. Розглянемо елемент $a_{ij}^{(m)}$ матриці A^m. Оскільки $A^m = A^{m-1}A$, то

$$a_{ij}^{(m)} = \sum_{t=1}^{n} a_{it}^{(m-1)} a_{tj} = \sum_{t=1}^{n} a_{it}^{(m-1)} a_{tj}^{(1)}.$$

Розглянемо окремий доданок $a_{it}^{(m-1)} a_{tj}^{(1)}$ останньої суми. За припущенням індукції $a_{it}^{(m-1)}$ дорівнює кількості шляхів довжини $m - 1$, які ведуть з вершини v_i у вершину v_t; тоді добуток $a_{it}^{(m-1)} a_{tj}^{(1)}$ дорівнює кількості шляхів довжини m, що ведуть із вершини v_i у вершину v_j і передостанньою вершиною яких є v_t. Отже, сума таких доданків для всіх t від 1 до n дає шукану кількість шляхів довжини m із v_i у v_j. Теорему доведено.

Наслідок 4.9.1. Нехай A — матриця суміжності графа $G = (V, E)$ з n вершинами. У графі G вершини v_i та v_j $(i \neq j)$ зв'язані тоді й тільки тоді, коли елемент i-го рядка та j-го стовпчика матриці $A + A^2 + A^3 + \ldots + A^{n-1}$ не дорівнює нулю.

Це випливає з доведеної теореми та тієї простої властивості, що коли у графі G з n вершинами існує шлях між вершинами v_i і v_j $(i \neq j)$, тоді між цими вершинами обов'язково існує шлях довжини не більшої ніж $n - 1$ (див. лему 4.2).

Крім того, щоб вилучити умову $i \neq j$ для виявлення зв'язності між будь-якими вершинами графа G, можна використовувати матрицю $M^{(n)} = I_n + A + A^2 + A^3 + \ldots + A^{n-1}$, де I_n — одинична матриця порядку n (нагадаємо, що будь-яка вершина зв'язана сама з собою шляхом довжини 0).

Наслідок 4.9.2. Граф G зв'язний тоді й тільки тоді, коли в матриці $M^{(n)}$ немає нульових елементів.

Граф $G^* = (V, E^*)$ називається ***транзитивним замиканням*** графа $G = (V, E)$, якщо $(v, w) \in E^*$ тоді й тільки тоді, коли вершини v і w зв'язані у графі G.

Отже, транзитивне замикання графа G є повним графом тоді й тільки тоді, коли граф G зв'язний.

Якщо графу $G = (V, E)$ відповідає відношення R на V, то графу G^* відповідатиме транзитивне замикання R^* відношення R.

Побудуємо для графа G^* $n \times n$-матрицю A^* за таким правилом: (i, j)-й елемент матриці A^* дорівнює 1 тоді й тільки тоді, коли відповідний елемент матриці $M^{(n)}$ не дорівнює 0, усі інші елементи матриці A^* дорівнюють 0.

Матрицю A^* називають ***матрицею досяжності*** графа G (інші назви: *матриця зв'язності, матриця зв'язку*).

Матрицю досяжності A^* можна обчислити й іншим методом.

Позначимо через $A^{(1)}$ булеву матрицю, елементи якої повністю збігаються з елементами матриці A, але їх розглядають не як числа 0 і 1, а як символи булевого алфавіту 0 і 1. Уведемо операцію булевого множення $C \wedge D$ матриць C і D, які складаються з булевих елементів 0 і 1, так: (i, j)-й елемент матриці $C \wedge D$ дорівнює $f_{ij} = \bigvee_{t=1}^{n}(c_{it} \wedge d_{tj})$, де c_{it} та d_{tj} — відповідні елементи матриць C і D, а операції $\vee$ і $\wedge$ — це операції диз'юнкції та кон'юнкції (див. розд. 2.3).

Позначимо через $A^{(m)}$ матрицю $A^{(1)} \wedge A^{(1)} \wedge \ldots \wedge A^{(1)}$ (m разів).

Аналогічно теоремі 4.9 можна довести таку теорему.

Теорема 4.10. Нехай $A^{(1)}$ — булева матриця, яка відповідає матриці суміжності A графа $G = (V, E)$. Елемент $b_{ij}^{(m)}$ $(i \neq j)$ матриці $A^{(m)}$ дорівнює 1 тоді й тільки тоді, коли у графі G існує принаймні один шлях довжини m, що веде з вершини v_i у v_j.

Наслідок 4.10.1. Матрицю досяжності A^* графа G з n вершинами можна обчислити за формулою $A^* = I_n^{(1)} \vee A^{(1)} \vee A^{(2)} \vee \ldots \vee A^{(n-1)}$ (операція диз'юнкції виконується для матриць поелементно).

Наслідок 4.10.2. Граф G зв'язний тоді й тільки тоді, коли всі елементи його матриці досяжності A^* дорівнюють 1.

4.5. Задачі і вправи

1. Користуючись наведеними алгоритмами, з'ясувати, чи є зв'язними графи, задані матрицями суміжності в задачі 4.1.2.

2. Нехай A — матриця суміжності графа G. Довести, що:

(а) діагональний елемент $a_{ii}^{(3)}$ матриці A^3 дорівнює подвоєній кількості трикутників графа G, які містять вершину з номером i;

(б) сума всіх діагональних елементів $a_{ii}^{(3)}$ матриці A^3 в шість разів перевищує кількість трикутників у графі G.

3. Нехай A — матриця суміжності зв'язного графа G. Довести, що відстань $d(v_i, v_j)$ між вершинами v_i і v_j $(i \neq j)$ дорівнює найменшому натуральному числу k, для якого елемент $a_{ij}^{(k)}$ i-го рядка та j-го стовпчика матриці A^k відмінний від 0.

4.7. Пошук шляхів у графі

Користуючись методами, викладеними в попередньому розділі, можна визначити зв'язність (тобто наявність або відсутність шляху) між будь-якими вершинами v_i і v_j заданого графа $G = (V, E)$. Однак часто буває потрібно не просто виявити існування шляху між заданими вершинами, але й знайти послідовність вершин і ребер (шлях), що веде з v_i у v_j. Крім того, для різних практичних застосувань теорії графів важливою є проблема систематичного обходу (перебору) всіх вершин і/або ребер графа.

Відомі два класичні методи розв'язання цих проблем: *алгоритми пошуку* (*обходу графа*) *вшир* і *вглиб*. Сформулюємо постановку проблеми пошуку й обидва методи її розв'язання.

Нехай задано граф $G = (V, E)$, $V_п \subseteq V$ — множину початкових вершин і $V_к \subseteq V$ — множину кінцевих (заключних, цільових) вершин графа G. Потрібно знайти шлях із деякої вершини $v \in V_п$ в одну з вершин $w \in V_к$, тобто знайти таку послідовність ребер $(v_i, v_{i+1}) \in E$, $i = 1, 2, \ldots, t-1$, що $v_1 = v$ і $v_t = w$.

Зокрема, множини початкових або кінцевих вершин можуть містити тільки по одній вершині. Такі вершини природно називати початковою та кінцевою вершинами графа G.

Для опису алгоритмів нам знадобляться три списки ребер: ВІДКР, ЗАКР і РОЗВ. Крім того, для вершини $v \in V$ через $S(v)$ будемо позначати множину всіх таких вершин $w \in V$, що у графі G існує ребро $(v, w) \in E$. Такі вершини w часто називають *синами* вершини v, а множину $S(v)$ — множиною синів вершини v.

Для зручності додамо до множини вершин V графа G "порожню" вершину p, а до множини ребер — "порожні" ребра виду (p, v), де $v \in V_п$. При визначенні шляху з $V_п$ у $V_к$ "порожні" ребра не враховують.

Оскільки в запропонованій нижче формі обидва алгоритми пошуку відрізняються тільки в одній позиції, викладемо їх одночасно і позначимо те місце, яким вони різняться між собою.

Алгоритм пошуку вшир (вглиб)

1. Усі "порожні" ребра розмістити у списку ВІДКР (у довільному порядку).

2. Якщо ВІДКР = $\varnothing$, то РОЗВ = $\varnothing$ (тобто сформульована задача не має розв'язку) й алгоритм завершує свою роботу.

3. Закрити перше ребро (v, w) з ВІДКР, тобто перенести ребро (v, w) зі списку ВІДКР у список ЗАКР.

4. Якщо вершина w закритого ребра кінцева ($w \in V_к$), то шуканий список РОЗВ (тобто шуканий шлях із $V_п$ у $V_к$) міститься серед ребер списку ЗАКР, і його можна послідовно виділити з цього списку, починаючи з останнього закритого ребра шляху. Зауважимо, що, будуючи шуканий шлях для кожного з ребер (v, w) списку ЗАКР, слід вибирати ребро-попередника (u, v) так, щоб воно було першим ребром із кінцем v в списку ЗАКР. Алгоритм завершує свою роботу.

В іншому випадку ($w \notin V_к$) перейти до пункту 5.

5. Визначити $S(w)$ — множину синів вершини w останнього закритого ребра, а також множину ребер $R(w) = \{(w, z) \mid z \in S(w)\}$.

Розмістити у списку ВІДКР усі ребра з множини $R(w) \setminus$ $\setminus$(ВІДКР $\cup$ ЗАКР) *після* всіх ребер (*перед* усіма ребрами), що вже містяться в цьому списку.

6. Перейти до пункту 2.

Ці алгоритми пошуку різняться лише в позиції 5, і відмінність між ними полягає в тому, що в алгоритмі пошуку вшир потрібно розміщувати відповідні ребра *після*, а в алгоритмі пошуку вглиб — *перед* усіма ребрами, що входять до списку ВІДКР.

Для наведених алгоритмів використовують скорочені назви **АПШ** й **АПГ** (відповідні англійські назви — BFS (breadth first search) і DFS (depth first search)).

Таким чином, для АПШ список ВІДКР є *чергою*, тобто такою сукупністю елементів, у якій нові елементи розміщують у кінці сукупності, а елемент, що "обслуговується" (закривається), вибирають із голови (початку) цієї сукупності. Водночас для АПГ список ВІДКР є так званим *стеком*, тобто сукупністю, у якій елементи, що додаються до сукупності і елементи, що відбираються для "обслуговування", розміщуються тільки на початку сукупності — у верхівці стеку (за принципом "останній прийшов — перший обслуговується").

Приклад 4.8. Розглянемо дію алгоритму пошуку вшир для графа, зображеного на рис. 4.6 (табл. 4.1). Вважаємо, що $V_п = \{3\}$ і $V_к = \{11\}$.

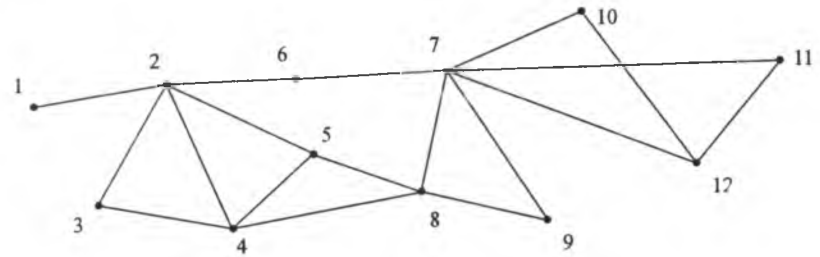


Рис. 4.6

Таблиця 4.1

Крок	ВІДКР	ЗАКР	*w*	*R*(*w*)
0	(*р*, 3)			
1	(3, 2), (3, 4)	(*р*, 3)	3	(3, 2), (3, 4)
2	(3, 4), (2, 1), (2, 4), (2, 5), (2, 6)	(3, 2)	2	(2, 1), (2, 4), (2, 5), (2, 6)
3	(2, 1), (2, 4), (2, 5), (2, 6), (4, 5), (4, 8)	(3, 4)	4	(4, 5), (4, 8)
4	(2, 4), (2, 5), (2, 6), (4, 5), (4, 8)	(2, 1)	1	
5	(2, 5), (2, 6), (4, 5), (4, 8)	(2, 4)	4	
6	(2, 6), (4, 5), (4, 8), (5, 8)	(2, 5)	5	(5, 8)
7	(4, 5), (4, 8), (5, 8), (6, 7)	(2, 6)	6	(6, 7)
8	(4, 8), (5, 8), (6, 7)	(4, 5)	5	
9	(5, 8), (6, 7), (8, 7), (8, 9)	(4, 8)	8	(8, 7), (8, 9)
10	(6, 7), (8, 7), (8, 9)	(5, 8)	8	
11	(8, 7), (8, 9)	(6, 7)	7	(7, 12), (7, 11), (7, 9), (7, 10)
12	(8, 9), (7, 12), (7, 11), (7, 9), (7, 10)	(8, 7)	7	
13	(7, 12), (7, 11), (7, 9), (7, 10)	(8, 9)	9	
14	(7, 11), (7, 9), (7, 10), (12, 10), (12, 11)	(7, 12)	12	(12, 10), (12, 11)
15	(7, 9), (7, 10), (12, 10), (12, 11)	(7, 11)	11	

Кожен рядок таблиці описує результат виконання одного циклічного кроку (позиції 2–6) алгоритму. Оскільки список ЗАКР на кожному кроці поповнюється тільки одним ребром, то в таблиці записуємо лише це ребро (не повторюючи всі елементи, які ввійшли до складу ЗАКР на попередніх кроках). Нагадаємо також, що ребра (v, w) і (w, v) збігаються.

Алгоритм завершує свою роботу на п'ятнадцятому кроці. Аналізуючи список ребер ЗАКР від його кінця до початку, будуємо список РОЗВ, тобто знаходимо шуканий шлях від вершини 3 у вершину 11: 3, (3, 2), 2, (2, 6), 6, (6, 7), 7, (7, 11), 11. Довжина цього шляху — 4.

Радимо побудувати аналогічну таблицю для алгоритму пошуку вглиб шляху з вершини 3 у вершину 11 і порівняти дію та результати обох алгоритмів.

4.8. Аналіз і модифікації алгоритмів пошуку

Відзначимо деякі найважливіші характеристики та порівняємо алгоритми пошуку між собою.

1. Обидва алгоритми гарантують знаходження розв'язку, якщо він існує. Це випливає з того, що обидва алгоритми аналізують усі вершини й усі ребра графа, тому або гарантовано знаходять шуканий шлях, або виявляють, що розв'язку немає.

2. Складність обох алгоритмів (тобто кількість кроків, які має виконати кожен алгоритм, щоб знайти розв'язок або виявити, що розв'язку не існує) не перевищує величини $O(n + m)$, де n — кількість вершин, а m — кількість ребер графа G.

Справді, кожне з m ребер графа аналізують не більше одного разу, а кількість кроків для визначення множини ребер $R(w)$ (крок 5 алгоритму) має порядок числа вершин графа.

3. Якщо $|V_п| = 1$, то АПШ знаходить шлях мінімальної довжини з $V_п$ у $V_к$. Якщо ж $|V_п| > 1$, то для пошуку найкоротшого шляху з $V_п$ у $V_к$ потрібно дещо модифікувати АПШ, а саме — знайти всі шляхи з вершин множини $V_п$ в множину вершин $V_к$, порівняти їх за довжиною й обрати мінімальний.

4. Розв'язок, який знаходить АПГ, узагалі кажучи, не оптимальний. У найгіршому випадку АПГ може знайти шлях із $V_п$ у $V_к$ довжини $n - 2$, хоча існує шлях довжини 1.

5. АПГ легше запрограмувати з таких причин:
– унаслідок використання стеку;
– завдяки можливості запису АПГ у компактній і наочній рекурсивній формі.

6. АПГ — складова частина багатьох важливих алгоритмів для графів (побудова кістякового дерева, пошук зв'язних компонент, топологічне

сортування вершин орграфа, пошук розділювальних вершин, перевірка планарності графа тощо [16; 21; 24]).

7. Пам'ять, яку використовує АПГ, узагалі кажучи, менша за обсягом, ніж та, що потрібна АПШ.

8. Аналіз людського мислення свідчить, що людині більш притаманний АПГ при пошуках розв'язків різних задач або під час проведення досліджень (пошуки шляху в лабіринті, аналіз шахових комбінацій, розв'язування логічних задач, спелеологічні дослідження тощо).

Розглянемо деякі важливі *модифікації* алгоритмів пошуку.

Далі для зручності вважатимемо, що $|V_п| = 1$.

У великих графах (тобто у графах зі значною кількістю вершин і ребер) АПГ має одну істотну ваду, яка є наслідком довільності розміщення ребер з $R(w)$ у списку ВІДКР (крок 5). Зробивши на якомусь кроці "помилковий" ("неоптимальний") вибір продовження шляху, АПГ може пройти повз наявний коротший шлях і заглибитися в граф у пошуках значно довшого й гіршого шляху.

Спробою усунути цю ваду є так званий ***алгоритм обмеженого пошуку вглиб*** (**АОПГ**). У ньому кожен шлях із $V_п$ простежують углиб доти, доки його довжина не досягне певного наперед заданого граничного значення k. Після цього алгоритм не поповнює список ВІДКР (крок 5), а починає досліджувати інший шлях, повернувшись до останнього розгалуження, і т. д.

На жаль, ще серйознішою вадою АОПГ порівняно з АПГ є те, що він може не знайти жодного розв'язку, хоча такий розв'язок є. Ця ситуація може виникнути тоді, коли довжина розв'язку більша від k.

Тому застосовують нижченаведену модифікацію, яка називається ***алгоритмом поступового*** (*прогресивного*) ***заглиблення*** (**АПЗ**).

Якщо для заданого k АОПГ не знайшов розв'язку, то значення k збільшується на певну величину t (тобто $k := k + t$), й АОПГ повторює свою роботу спочатку.

Коли $k \geq n - 1$, де $n = |V|$, то АПЗ збігається з АОПГ й обидва вони збігаються з класичним АПГ. Якщо ж $k = 1$ і $t = 1$, то АПЗ збігається з АПШ.

Таким чином, добираючи значення k і t, можна реалізувати певний компроміс між АПГ і АПШ.

Обидва алгоритми можна дуже просто модифікувати для того, щоб виконати систематичний обхід усіх вершин і/або всіх ребер графа. Результатом такого обходу є певна нумерація (порядок розміщення) вершин і/або ребер.

Цю модифікацію одержимо, якщо вилучимо у кроці 4 перевірку $w \in V_к$ (а отже, фактично вилучимо весь крок 4), і єдиною умовою завершення алгоритму пошуку залишимо умову кроку 2 (список ВІДКР — порожній).

Відповіддю (результатом дії) такого алгоритму буде або список ребер у ЗАКР, і/або та послідовність, у якій було проаналізовано вершини на кроці 5 (без повторень).

Нарешті, ще однією цікавою і важливою задачею, для розв'язання якої використовують наведені алгоритми, є задача пошуку шляху найменшої вартості.

Нехай задано граф $G = (V, E)$ і деяку дійсну функцію $f : E \to R^+$, де R^+ — множина невід'ємних дійсних чисел. Такий граф, кожному ребру якого поставлено у відповідність невід'ємне дійсне число, називане ***вартістю*** (*вагою*, *ціною*) ребра, називається ***зваженим*** (*позначеним*) графом. Тоді ***вартістю*** (*вагою*) ***шляху*** M, ребрами якого є $e_{i1}, e_{i2}, \ldots, e_{ik}$, називають величину $W = \sum_{j=1}^{k} f(e_{ij})$.

Нехай $V_п \subseteq V$ і $V_к \subseteq V$ — множини початкових і кінцевих вершин зваженого графа $G = (V, E)$. Алгоритми, які шукають шляхи найменшої вартості з $V_п$ у $V_к$, називаються *алгоритмами пошуку найкоротших шляхів*. Їх описано в [16; 21; 24].

4.6. Задачі і вправи

1. Написати алгоритм для реалізації обмеженого пошуку вглиб (АОПГ).

2. Написати алгоритм пошуку в графі, що реалізує метод поступового заглиблення (АПЗ).

3. Написати модифікації алгоритмів пошуку (углиб і вшир), які визначають деяке кістякове дерево даного графа (див. розділ 4.9).

4. Написати модифікацію алгоритму пошуку вглиб, яка визначає, чи має заданий граф цикл.

5. Написати алгоритм, який використовує стратегію пошуку вглиб і:

(а) здійснює систематичний обхід усіх вершин даного графа;

(б) визначає всі зв'язні компоненти;

(в) визначає зв'язність графа;

(г) будує матрицю досяжності графа;

(д) знаходить усі цикли графа.

6. Написати алгоритм, який використовує стратегію пошуку вшир і:

(а) здійснює систематичний обхід усіх вершин даного графа;

(б) визначає всі зв'язні компоненти;

(в) визначає зв'язність графа;

(г) будує матрицю досяжності графа;

(д) визначає відстані між усіма парами вершин;

(е) визначає ексцентриситети всіх вершин;

(є) обчислює радіус і діаметр графа;

(ж) визначає центр графа.

4.9. Деякі важливі класи графів: дерева та двочасткові графи

Граф без циклів називається ***ациклічним***, ациклічний зв'язний граф — ***деревом***, довільний ациклічний граф — ***лісом***.

Зв'язними компонентами лісу є дерева, тому кожен ліс можна зобразити у вигляді прямої суми дерев.

Дерева — це особливий і дуже важливий клас графів, бо, по-перше, їх широко застосовують у різних галузях науки і практики, а, по-друге, вони займають особливе положення у самій теорії графів. Останнє випливає з граничної простоти будови дерев. Часто, розв'язуючи різні задачі теорії графів, їх дослідження починають із дерев. Зокрема, порівняно нескладною є проблема перевірки ізоморфності дерев.

Існують й інші, рівносильні наведеному, означення дерева, які можна розглядати як характеристичні властивості дерева.

Теорема 4.11. Для графа $G=(V, E)$, $|V| = n$, $|E| = m$, такі твердження рівносильні:

1) G — дерево (ациклічний зв'язний граф);

2) G — зв'язний граф, і $m = n - 1$;

3) G — ациклічний граф, і $m = n - 1$;

4) для будь-яких вершин v і w графа G існує лише один простий ланцюг, що їх з'єднує;

5) G — такий ациклічний граф, що коли будь-які його несуміжні вершини v і w з'єднати ребром (v, w), то одержаний граф міститиме рівно один цикл.

Доведення. Для доведення теореми покажемо виконання такого ланцюжка логічних слідувань: 1) $\Rightarrow$ 2), 2) $\Rightarrow$ 3), 3) $\Rightarrow$ 4), 4) $\Rightarrow$ 5) і 5) $\Rightarrow$ 1).

Оскільки відношення логічного слідування транзитивне, то звідси випливатиме рівносильність усіх п'яти тверджень.

Для тривіального графа G ($n = 1$) справедливість твердження теореми очевидна, тому вважатимемо, що $n > 1$.

1) $\Rightarrow$ 2). Доведемо це твердження методом математичної індукції за значенням n. Для $n = 2$ умову 1) задовольняє тільки один граф K_2, він же задовольняє й умову 2).

Припустімо, що твердження виконується для всіх дерев із кількістю вершин $n \leq t$ ($t \geq 2$). Розглянемо довільне дерево $G = (V, E)$, в якому $t + 1$ вершина. Вилучимо з G деяке ребро $e \in E$. За теоремою 4.7, *б* отримаємо граф G', що складається з двох ациклічних зв'язних компонент, тобто з двох дерев T_1 і T_2. Нехай дерево T_1 має n_1 вершин і m_1 ребер, а дерево T_2 — n_2 вершин і m_2 ребер, $n_1 \leq t$ і $n_2 \leq t$. За припущенням індукції, маємо $m_1 = n_1 - 1$ і $m_2 = n_2 - 1$. Отже, для зв'язного графа G виконуються такі рівності: $m = m_1 + m_2 + 1 = (n_1 - 1) + (n_2 - 1) + 1 = n_1 + n_2 - 1 = (t + 1) - 1 = t$.

2) $\Rightarrow$ 3). Доведемо методом від супротивного. Припустімо, що у графі G є цикл. Вилучивши в G довільне ребро e цього циклу, за теоремою 4.7, *а* дістанемо зв'язний граф G', у якому $n - 2$ ребра. Останнє суперечить наслідку 4.8.1. Отже, граф G ациклічний.

3) $\Rightarrow$ 4). Знову скористаємося методом доведення від супротивного. Припустімо, що для графа G виконується умова 3), але граф G незв'язний і має k компонент зв'язності. Тоді кожна з цих зв'язних компонент T_i ациклічна, тобто є деревом. Нехай дерево T_i має n_i вершин і m_i ребер, $i = 1, 2, \ldots, k$. Із доведеного вище маємо $m_i = n_i - 1$, $i = 1, 2, \ldots, k$. Тоді $n - 1 = m = m_1 + m_2 + \ldots + m_k = (n_1 - 1) + (n_2 - 1) + \ldots + (n_k - 1) = (n_1 + n_2 + \ldots + n_k) - k = n - k$.

Отже, $k = 1$, і G — зв'язний граф.

Відтак припустімо, що граф G задовольняє умову 3), але має дві вершини v і w, які можна з'єднати двома різними простими ланцюгами. Ці ланцюги утворюють циклічний маршрут, що веде з v у v й обов'язково містить у собі деякий цикл (доведіть це самостійно). Останнє суперечить умові 3).

4) $\Rightarrow$ 5). Якщо припустити, що у графі G є цикл, тоді будь-які дві вершини цього циклу можна з'єднати між собою принаймні двома простими ланцюгами. Отже, G — ациклічний граф. Візьмемо будь-які дві несуміжні вершини v і w у графі G й додамо до нього ребро (v, w); дістанемо граф G'. У ньому є один цикл Z, який складається з просто-

го ланцюга, що веде з v у w в графі G, та доданого ребра (v, w). Припустімо, що у графі G' є ще один цикл Z_1 ($Z_1 \neq Z$). Цикли Z_1 і Z мають спільні ребра (в іншому випадку Z_1 є циклом ациклічного графа G). Якщо серед цих ребер немає ребра (v, w), то знову отримаємо Z_1 — цикл у графі G. Отже, цикли Z і Z_1 мають спільне додане ребро (v, w). Тоді частина циклу Z, що веде із v у w, разом із частиною циклу Z_1, що веде з w у v, утворює замкнений (циклічний) маршрут, що веде з v у v в графі G. Зазначені частини циклів Z і Z_1 не збігаються, тому цей циклічний маршрут містить у собі цикл, що суперечить ациклічності графа G.

5) $\Rightarrow$ 1). Необхідно довести, що G — зв'язний граф. Припустімо, що це не так. Візьмемо дві довільні вершини v і w з двох різних компонент зв'язності графа G та з'єднаємо їх ребром; дістанемо граф G'. Оскільки обидві компоненти є ациклічними графами, то граф G' також не міститиме циклів. Це суперечить умові 5).

Теорему 4.11 доведено.

Наслідок 4.11.1. Для довільного дерева $T = (V, E)$ з n вершинами виконується $\sum_{v\in V}\delta(v) = 2(n-1)$.

Наслідок 4.11.2. Будь-яке нетривіальне дерево $T = (V, E)$ має принаймні дві кінцеві вершини.

Припустімо, що дерево T має менше двох кінцевих вершин. Тоді степінь лише однієї вершини може дорівнювати 1, а степені всіх інших не менші 2. Отже, $\sum_{v\in V}\delta(v) \geq 2(n-1)+1 = 2n-1$, що суперечить наслідку 4.11.1.

Наслідок 4.11.3. Ліс F, який має n вершин і складається з k дерев, містить $n - k$ ребер.

Справді, якщо дерево T_i лісу F має n_i вершин, то за доведеною теоремою воно містить $n_i - 1$ ребро, $i = 1, 2, \ldots, k$. Додаючи кількості ребер кожного з дерев T_i, дістанемо кількість $n - k$ ребер у F.

Наслідок 4.11.4. У графі G з n вершинами, який має більше ніж $n - 1$ ребро, є принаймні один цикл.

Розглянемо довільний граф G з n вершинами та кількістю ребер, яка перевищує $n - 1$. Припустімо, що G — ациклічний граф. Тоді G — ліс, що складається з k дерев ($k \geq 1$). За попереднім наслідком кількість ребер у такому графі дорівнює $n - k$; тоді $n - k > n - 1$, тобто $k < 1$, що неможливо.

Кістяковим (*каркасним*) ***деревом*** зв'язного графа $G = (V, E)$ називається дерево $T = (V, E_T)$ таке, що $E_T \subseteq E$.

Кістяковим (*каркасним*) ***лісом*** незв'язного графа $G = (V, E)$ називається сукупність кістякових (каркасних) дерев зв'язних компонент графа G.

Наслідок 4.11.5. Для зв'язного графа $G = (V, E)$ можна зазначити $|E| - |V| + 1$ ребро, після вилучення яких отримаємо кістякове дерево графа G.

Із теореми 4.7 випливає, що потрібно послідовно вилучати ребра, які належать циклам [16; 21; 24]. Порядок вилучення неістотний. Кількість ребер, що залишаться в кістяковому дереві графа G, дорівнює $|V| - 1$; отже, має бути вилучено $|E| - (|V| - 1) = |E| - |V| + 1$ ребро.

Наслідок 4.11.6. Нехай граф $G = (V, E)$ має k компонент зв'язності. Для отримання його кістякового лісу з графа G потрібно вилучити $|E| - |V| + k$ ребер.

Для доведення цього твердження потрібно застосувати попередній наслідок до кожної компоненти зв'язності графа G, відтак підсумувати результати.

Число $|E| - |V| + k$ називають ***цикломатичним числом*** графа G і позначають $\nu(G)$.

Пропонуємо самостійно довести такі прості властивості цикломатичного числа $\nu(G)$ графа G.

Лема 4.3. 1. Для довільного графа G виконується $\nu(G) \geq 0$.

2. Граф G є лісом тоді й тільки тоді, коли $\nu(G) = 0$.

3. Граф G має рівно один простий цикл тоді й тільки тоді, коли $\nu(G) = 1$.

4. Кількість циклів у графі G не менша ніж $\nu(G)$.

Алгоритми знаходження кістякових дерев (кістякових лісів) для заданих графів можна побудувати на основі вищезгаданих алгоритмів пошуку вшир або вглиб [16; 21; 24].

Граф $G = (V, E)$ називається ***двочастковим***, якщо існує таке розбиття $\{V_1, V_2\}$ множини його вершин V на дві підмножини (*частки*), що для довільного ребра $(v, w) \in E$ або $v \in V_1$ і $w \in V_2$, або $v \in V_2$ й $w \in V_1$.

Двочастковий граф $G = (V, E)$ називається ***повним двочастковим***, якщо для будь-якої пари вершин його часток $v \in V_1$ і $w \in V_2$ маємо $(v, w) \in E$. Якщо $|V_1| = m$ і $|V_2| = n$, то повний двочастковий граф G позначають $K_{m,n}$.

Теорема 4.12 (Кеніга). Граф є двочастковим тоді й тільки тоді, коли всі його цикли мають парну довжину.

Доведення. Необхідність. Нехай G — двочастковий граф, а Z — його цикл. Будь-який маршрут у графі G, що веде з довільної вершини v однієї частки V' у будь-яку вершину тієї самої частки, завжди має парну довжину, бо всі непарні ребра цього маршруту ведуть із частки V' в іншу частку, а всі парні ребра, навпаки, повертають маршрут у V'. Отже, і довжина циклу Z — парне число.

Достатність. Нехай усі цикли графа G мають парну довжину. Розглянемо довільну зв'язну компоненту $G' = (V', E')$ графа G. Візьмемо вершину v цієї компоненти й побудуємо розбиття $\{V_1, V_2\}$ множини вершин V' на дві частки таким чином: віднесемо до V_1 усі вершини $w \in V'$, для яких відстань $d(v, w)$ — парне число, а до V_2 — усі вершини $u \in V'$, для яких відстань $d(v, u)$ — непарне число.

Доведемо, що жодні дві вершини з однієї частки несуміжні у графі G. Припустімо, шо v_1 і v_2 — дві різні вершини з однієї частки, й існує ребро $(v_1, v_2) \in E'$. Тоді жодна з вершин v_1 і v_2 не збігається з v, оскільки $v \in V_1$, а всі вершини, суміжні з v, належать частці V_2.

Позначимо через L_1 і L_2 найкоротші прості ланцюги, що ведуть із v відповідно у v_1 і v_2. Довжини цих ланцюгів мають однакову парність, бо, за припущенням, v_1 і v_2 належать одній частці. Нехай u — остання спільна вершина L_1 і L_2 (починаючи від v). Довжини частин обох цих ланцюгів, що ведуть з u у v_1 і з u у v_2, також матимуть однакові парності. Тоді маршрут, який складається з частки L_1, що веде з u у v_1, містить ребро (v_1, v_2) й завершується часткою L_2, що веде з v_2 в u (тобто його проходять у зворотному порядку), є циклом, довжина якого непарна. Це суперечить умові. Отже, будь-яка зв'язна компонента графа G є двочастковим графом, а тому й сам граф G двочастковий.

Наслідок 4.12.1. Граф є двочастковим тоді й тільки тоді, коли він не має простих циклів непарної довжини.

Наслідок 4.12.2. Будь-яке дерево — двочастковий граф.

Наслідок 4.12.3. Простий цикл парної довжини C_{2k} — двочастковий граф.

4.7. Задачі і вправи

1. Чи може зв'язний граф з n вершинами та $n - 1$ ребром мати цикл? Відповідь об'рунтувати.

2. Довести, що кількість кінцевих вершин у дереві з n вершинами, серед яких немає вершин степеня 2, не менша від $n/2 + 1$.

3. Нехай у графі G з n вершинами ($n \geq 3$) кількість кінцевих вершин збігається з кількістю ребер. Довести, що граф G або є незв'язним, або є деревом.

4. Описати всі дерева, доповнення яких є також деревами.

5. Описати всі дерева, які є самодоповнювальними графами.

6. Довести, що граф G зв'язний тоді й тільки тоді, коли він має кістякове дерево.

7. Довести, що цикломатичне число $\nu(G)$ довільного графа G невід'ємне.

8. Нехай $\nu(G)$ — цикломатичне число графа G. Довести, що:
(а) граф G є лісом тоді й тільки тоді, коли $\nu(G) = 0$;
(б) граф G має тільки один простий цикл тоді й тільки тоді, коли $\nu(G) = 1$;
(в) кількість циклів у графі G не менша ніж $\nu(G)$.

9. Скільки ребер містить повний двочастковий граф $K_{n,m}$?

10. Довести, що будь-яке дерево є двочастковим графом. Які дерева є повними двочастковими графами?

11. Який вигляд має доповнення графа $K_{n,m}$?

12. Чи для кожного натурального числа k існує повний двочастковий граф, кількість ребер якого дорівнює k ($k \geq 2$)?

13. Визначити для зв'язного двочасткового графа з n вершинами найменшу і найбільшу можливу кількість ребер.

14. Довести, що граф G з n вершинами не є двочастковим, якщо кількість його ребер більша від $n^2/4$.

15. Чому дорівнюють діаметр і радіус повного двочасткового графа $K_{n,m}$?

16. Нехай A — матриця суміжності графа G. Довести, що граф G двочастковий тоді й тільки тоді, коли для довільного непарного числа n усі діагональні елементи матриці A^n дорівнюють 0.

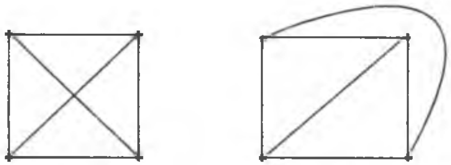
4.10. Плоскі та планарні графи

Часто не має особливого значення, як зобразити граф у вигляді рисунка на площині (діаграми), бо ізоморфні графи подібні за своєю структурою і містять ту саму інформацію. Однак існують ситуації, коли потрібно, щоб зображення графа на площині задовольняло певні умови. Наприклад, якщо граф є моделлю якоїсь електронної схеми чи транспортної мережі, де вершини позначають окремі елементи схеми чи станції, а ребра — відповідно електричні провідники або шляхи, то бажано так розмістити ці ребра на площині, щоб уникнути перетинів. Так виникає поняття плоского графа.

Граф називається ***плоским***, якщо його діаграму можна зобразити на площині так, що лінії, які відповідають ребрам графа, не перетинаються (тобто мають спільні точки тільки у вершинах графа). Таке зображення називається ***плоскою картою*** графа.

Граф називають ***планарним***, якщо він ізоморфний деякому плоскому графу.

Наприклад, граф, зображений на рис. 4.7, *а*, планарний, бо він ізоморфний графу, зображеному поруч (рис. 4.7, *б*). Простий цикл, дерево і ліс — також планарні графи.



а *б*

Рис. 4.7

Неважко обґрунтувати такі твердження.

Лема 4.4. 1). Будь-який підграф планарного графа планарний.

2). Граф є планарним тоді й тільки тоді, коли кожна його зв'язна компонента — планарний граф.

Про планарні графи кажуть, що вони *укладаються на площині*, або мають *плоске укладання*.

Жордановою кривою назвемо неперервну лінію на площині, яка не перетинає сама себе.

Гранню плоского графа називається множина точок площини, кожну пару яких можна з'єднати жордановою кривою, що не перетинає ребер графа. ***Межею грані*** вважають замкнений маршрут, що обмежує цю грань.

Як приклад на рис. 4.8 зображено плоский граф із п'ятьма гранями.

Отже, плоский граф розбиває всю множину точок площини на грані так, що кожна точка належить деякій грані. Зазначимо, що плоский граф має одну, до того ж єдину, необмежену грань (на рис. 4.8 це грань *5*). Будемо називати її *зовнішньою*, а всі інші грані — *внутрішніми*.

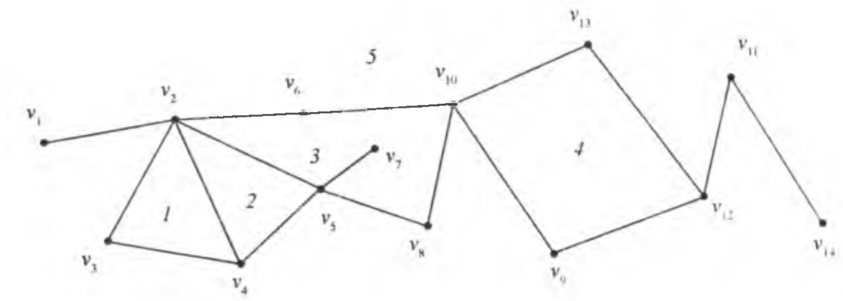


Рис. 4.8

Множину граней плоского графа позначатимемо через P.

Степенем грані r називають довжину циклічного шляху, що обмежує грань r (тобто довжину межі грані r); позначається Δ_r.

Для плоского графа на рис. 4.8 $\{v_3, (v_3, v_2), v_2, (v_2, v_4), v_4, (v_4, v_3), v_3\}$ — циклічний шлях для грані 1, $\{v_2, (v_2, v_5), v_5, (v_5, v_7), v_7, (v_7, v_5), v_5, (v_5, v_8), v_8, (v_8, v_{10}), v_{10}, (v_{10}, v_6), v_6, (v_6, v_2), v_2\}$ — циклічний шлях для грані 3. Отже, $\Delta_1 = 3$ та $\Delta_3 = 7$.

Лема 4.5. Нехай $G = (V, E)$ — плоский граф. Тоді $\sum_{r\in P} \Delta_r = 2|E|$.

Доведення. Справді, кожне ребро плоского графа або розділяє дві різні грані, або лежить усередині однієї грані. Отже, кожне ребро графа G або входить у межі тільки двох граней, або є елементом межі лише однієї грані, але при циклічному обході цієї грані таке ребро проходять двічі. Тому кожне ребро плоского графа вносить у розглядувану суму дві одиниці.

Теорема 4.13 (Ейлера). Для будь-якого зв'язного плоского графа $G = (V, E)$ виконується рівність

$$|V| - |E| + |P| = 2. \qquad (4.3)$$

Доведення. Нехай $G = (V, E)$ — зв'язний плоский граф з $n = |V|$ вершинами, а $T = (V, E_T)$ — деяке його кістякове дерево. Дерево T має тільки одну грань (зовнішню). Кількість ребер дерева T дорівнює $|E_T| = |V| - 1$. Отже, для кістякового дерева T формула (4.3) виконується.

Відтак будемо послідовно проводити в дереві T ребра графа G з множини $E \setminus E_T$. При цьому на кожному кроці цієї процедури кількість вершин $|V|$ залишатиметься незмінною, а кількість ребер і граней (див. теорему 4.11) одночасно збільшуватимуться на одиницю. Таким чином, формула Ейлера (4.3) виконується після кожної такої операції, тому вона справджується й для графа G, який отримаємо на завершення всієї процедури.

Наслідок 4.13.1. Кількість граней будь-якого плоского укладання зв'язного планарного графа з n вершинами й m ребрами є величиною сталою і дорівнює $m - n + 2$, тобто $|P| = |E| - |V| + 2$.

Інакше кажучи, число $|P|$ — це інваріант для заданого планарного графа G, тобто воно не залежить від способу укладання його на площині.

Наслідок 4.13.2. Для довільного зв'язного планарного графа $G = (V, E)$ з не менше ніж трьома вершинами виконується нерівність $|E| \le 3|V| - 6$.

Оскільки у графі G немає петель і кратних ребер, то степінь Δ_r будь-якої грані не менша 3, тобто $\sum_{r \in P} \Delta_r \ge 3|P|$.

Звідси, враховуючи співвідношення з леми 4.5 і попереднього наслідку, маємо $\sum_{r \in P} \Delta_r = 2|E| \ge 3(|E| - |V| + 2)$ та, нарешті, $|E| \le 3|V| - 6$.

Наслідок 4.13.3. У будь-якому планарному графі є принаймні одна вершина, степінь якої не перевищує 5.

Справді, якщо припустити, що степені всіх вершин планарного графа $G = (V, E)$ більші, ніж 5, то дістанемо нерівність $2|E| = \sum_{v \in V} \delta(v) \ge 6|V|$, яка суперечить попередньому наслідку.

Формулу Ейлера можна узагальнити.

Наслідок 4.13.4. Для довільного плоского графа $G = (V, E)$ з k компонентами зв'язності виконується рівність

$$|V| - |E| + |P| = k + 1. \tag{4.4}$$

Для доведення узагальненої формули Ейлера можна побудувати кістяковий ліс F графа G й переконатись у тому, що наведена рівність для F виконується (див. наслідок 4.11.3), відтак повторити міркування теореми 4.13.

Цей наслідок можна довести й іншим методом.

Якщо V_i, E_i та P_i — відповідно множини вершин, ребер і граней i-ї зв'язної компоненти графа G, то $|P_i| = |E_i| - |V_i| + 2$. Різні компоненти не мають спільних внутрішніх граней, а зовнішня грань для всіх компонент єдина, і її рахують один раз для кожної з них. Тому загальна кількість граней графа G дорівнюватиме

$$|P| = \sum_{i=1}^{k} |P_i| - (k-1) = \sum_{i=1}^{k} \left(|E_i| - |V_i| + 2\right) - (k-1) = \sum_{i=1}^{k} |E_i| - \sum_{i=1}^{k} |V_i| + 2k - (k-1) =$$
$$= |E| - |V| + k + 1.$$

Звідси дістанемо формулу (4.4).

Наслідок 4.13.5. Кількість внутрішніх граней довільного плоского графа G дорівнює цикломатичному числу $\nu(G)$ графа G.

Щоб переконатись у справедливості цього твердження, потрібно порівняти означення цикломатичного числа $\nu(G)$ (див. розд. 4.9) та узагальнену формулу Ейлера (4.4), узявши до уваги, що кількість внутрішніх граней дорівнює $|P| - 1$.

Наслідок 4.13.6. Графи G і G не можуть бути одночасно планарними, якщо кількість вершин у них не менша 11.

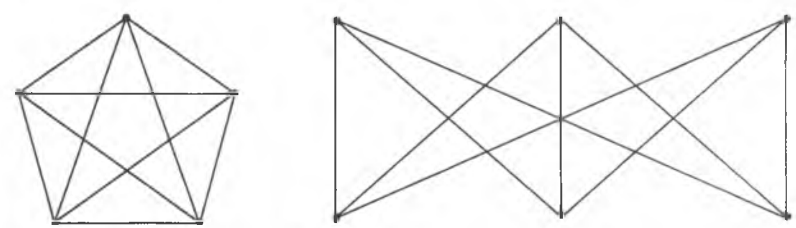
Справедливість твердження випливає з того, що нерівність із наслідку 4.13.2 не може одночасно виконуватися для графів G та G з кількістю вершин $|V| \ge 11$.

Максимальним планарним графом називається планарний граф, який при додаванні до нього будь-якого ребра перестає бути планарним.

Плоский зв'язний граф, кожну грань якого (включаючи й зовнішню) обмежено трикутником, називається ***тріангуляцією***.

Можна довести, що граф є максимальним плоским графом тоді й тільки тоді, коли він — тріангуляція.

У дослідженні плоских графів особливе місце займають графи K_5 і $K_{3,3}$, зображені на рис. 4.9.



K_5 $K_{3,3}$

Рис. 4.9

Теорема 4.14. Графи K_5 і $K_{3,3}$ не є планарними.

Доведення. Доведемо, що граф K_5 непланарний. Припустімо супротивне, тобто що $K_5 = (V, E)$ — планарний граф. Тоді з наслідку 4.13.2

випливає, що $|E| \le 3|V| - 6$. Однак, для графа K_5 $|E| = 10$ і $|V| = 5$, тобто повинно виконуватись $10 \le 3 \cdot 5 - 6 = 9$, що неможливо. Отже, припущення про те, що K_5 — планарний граф, неправильне.

Аналогічно, методом від супротивного доведемо, що граф $K_{3,3}$ непланарний. У графі $K_{3,3}$ жодні три вершини не є вершинами трикутника. Отже, $\Delta_r \ge 4$ для всіх граней $r \in P$. Припускаючи, що граф $K_{3,3}$ планарний, із наслідку 4.13.1 отримаємо $|P| = |E| - |V| + 2 = 9 - 6 + 2 = 5$. Тоді $2|E| = \sum_{r \in P} \Delta_r \ge 4|P| = 4 \cdot 5 = 20$, тобто $|E| \ge 10$, що неправильно для графа $K_{3,3}$.

Теорему доведено.

Значення графів K_5 і $K_{3,3}$ полягає в тім, що вони є "єдиними" істотно непланарними графами. Усі інші непланарні графи містять у собі підграфи, "подібні" до K_5 або $K_{3,3}$. Характер цієї подібності розкривається за допомогою таких понять.

Нехай $e = (v, w)$ — ребро графа G, а u — не є вершиною G. Вилучимо ребро e з графа G і додамо до нього нові ребра $e_1 = (v, u)$ і $e_2 = (w, u)$. Цю операцію називатимемо ***підрозбиттям ребра*** e.

Графи називаються ***гомеоморфними***, якщо їх можна отримати з одного графа за допомогою послідовного підрозбиття його ребер.

Приклад 4.9. На рис. 4.10 зображено два гомеоморфні графи G та G'.

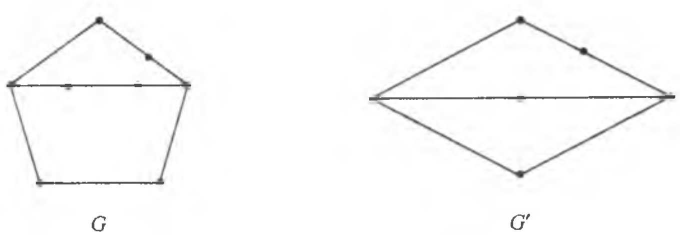


Рис. 4.10

Якщо граф G планарний, то будь-який граф, гомеоморфний G, також планарний.

Наведемо без доведення важливу теорему теорії графів [21, 33].

Теорема 4.15 (***Куратовського***). Граф G планарний тоді й тільки тоді, коли він не містить підграфів, гомеоморфних K_5 або $K_{3,3}$.

У теорії графів існують й інші критерії планарності. Наведемо ще один з них.

Елементарним стягуванням графа $G = (V, E)$ називається вилучення в ньому якогось ребра $(v_i, v_j) \in E$ та злиття вершин v_i і v_j в одну вершину v, причому v інцидентна всім тим відмінним від (v_i, v_j) ребрам графа G, які були інцидентні або вершині v_i, або вершині v_j.

Кажуть, що граф G ***стягується*** до графа G', якщо G' можна отримати з G за допомогою послідовності елементарних стягувань.

Приклад 4.10. Граф G з рис. 4.10 стягується до графа G', зображеного поруч.

Виходячи з теорем 4.14 і 4.15, неважко обґрунтувати таке твердження: *граф планарний тоді й тільки тоді, коли він не містить підграфів, що стягуються до* K_5 *або* $K_{3,3}$.

Наведені критерії було покладено в основу перших алгоритмів перевірки планарності графів. Однак час роботи таких алгоритмів пропорційний $|V|^6$. Згодом було створено значно швидші алгоритми, що здійснювали перевірку планарності графа $G = (V, E)$ за час $O(|V|)$ [24]. У монографіях [21; 24] описано також процедури, які для планарних графів знаходять плоскі укладання.

4.8. Задачі і вправи

1. Довести, що будь-який підграф планарного графа планарний.

2. Скільком граням може належати вершина степеня k плоского графа?

3. Довести, що будь-яке дерево є планарним графом. Скільки граней має дерево?

4. Чому дорівнює степінь єдиної грані дерева з n вершинами?

5. Знайти зв'язний плоский граф з n вершинами і m ребрами, для якого $m > 3n - 6$.

6. Чи існує планарний граф, який має:

(а) 7 вершин і 16 ребер;

(б) 8 вершин і 18 ребер?

7. Чи існує плоский граф із шістьма вершинами, що має дев'ять граней?

8. Побудувати планарний граф із вісьмома вершинами, доповнення якого є планарним графом.

9. Довести, що для зв'язного плоского графа з n вершинами ($n \ge 3$) й m ребрами, який не містить трикутників, виконується нерівність $m \le 2n - 4$.

10. Довести, що в будь-якому плоскому графі є або вершина степеня менше 3, або грань степеня менше 6.

11. Довести, що граф є максимальним плоским графом тоді й тільки тоді, коли він є тріангуляцією.

12. Довести, що будь-яка тріангуляція з n вершинами ($n \geq 3$) містить $3n - 6$ ребер і має $2n - 4$ граней.

4.11. Розфарбування графів

Нехай $G = (V, E)$ — довільний граф, а $N_k = \{1, 2, \ldots, k\}$.

Будь-яке відображення $f : V \to N_k$, яке кожній вершині $v \in V$ ставить у відповідність деяке натуральне число $f(v) \in N_k$, називається ***розфарбуванням*** графа G. Число $f(v)$ називають ***кольором***, або ***номером фарби***, вершини v.

Розфарбування f графа G називається ***правильним***, якщо для будь-яких його суміжних вершин v та w виконується $f(v) \neq f(w)$. Мінімальне число k, для якого існує правильне розфарбування графа G, називається ***хроматичним числом*** графа G й позначається $\chi(G)$.

Мінімальним правильним розфарбуванням графа G називається правильне розфарбування для $k = \chi(G)$.

Для певних типів графів нескладно визначити хроматичні числа. Наприклад, 1-хроматичними є порожні графи $G = (V, \varnothing)$ і тільки вони. Хроматичне число повного графа K_n дорівнює n, а хроматичне число довільного двочасткового графа — 2; 2-хроматичні графи часто називають ***біхроматичними***.

Неважко довести такі твердження.

Лема 4.6. Якщо кожна зв'язна компонента графа G потребує для свого правильного розфарбування не більше k фарб, то $\chi(G) \leq k$.

Лема 4.7. Граф є біхроматичним тоді й тільки тоді, коли він двочастковий.

Зокрема, всі дерева і прості цикли парної довжини C_{2k} біхроматичні. Водночас $\chi(C_{2k+1}) = 3$.

Використовуючи теорему 4.12, останню лему можна переформулювати в такому вигляді.

Лема 4.8. Граф є біхроматичним тоді й тільки тоді, коли він не має циклів непарної довжини.

Проблема визначення, чи є заданий граф k-хроматичним для певного k, та проблема відшукання мінімального правильного розфарбування

для заданого графа належать до класу задач, для яких на сьогодні не існують (і є всі підстави вважати, що не існують взагалі) ефективні точні алгоритми їх розв'язання [16; 21; 24]. Тому важливими є результати, що дають змогу оцінити значення хроматичного числа $\chi(G)$, виходячи з певних характеристик і властивостей графа G.

Теорема 4.16. Позначимо через $\Delta(G)$ найбільший зі степенів вершин графа G, тоді $\chi(G) \leq \Delta(G) + 1$.

Доведення проведемо індукцією за кількістю n вершин графа G. Для тривіального графа ($n = 1$) і графів із двома вершинами нерівність виконується.

Нехай твердження теореми виконується для всіх графів із кількістю вершин t ($t \geq 2$). Розглянемо довільний граф G з $t + 1$ вершиною. Вилучимо з нього деяку вершину v. Дістанемо граф G', степені всіх вершин якого не перевищують $\Delta(G)$. Отже, за припущенням індукції, для правильного розфарбування G' потрібно не більше ніж $\Delta(G) + 1$ фарб. Правильне розфарбування для G дістанемо з правильного розфарбування графа G', якщо пофарбуємо вершину v в колір, відмінний від кольорів усіх суміжних із v вершин. Оскільки таких вершин не більше ніж $\Delta(G)$, то для правильного розфарбування графа G достатньо $\Delta(G) + 1$ фарб.

Наслідок 4.16.1. Для правильного розфарбування довільного кубічного графа достатньо чотири фарби.

Так склалося історично, що окреме місце в теорії графів займають дослідження з розфарбування планарних графів. Це пов'язано зі славетною ***проблемою (гіпотезою) чотирьох фарб***.

Грані плоскої карти назвемо *суміжними*, якщо їх межі мають принаймні одне спільне ребро.

Гіпотеза чотирьох фарб виникла у зв'язку з розфарбуванням друкованих географічних карт (звідси й термін "плоска карта") і була сформульована так: "*Грані довільної плоскої карти можна розфарбувати не більше ніж чотирма фарбами так, що будь-які суміжні грані матимуть різні кольори*".

Згодом з'явилось інше, рівносильне, формулювання гіпотези чотирьох фарб: *для правильного розфарбування вершин довільного планарного графа потрібно не більше чотирьох фарб*.

Ця гіпотеза виникла в середині XIX століття. Більше ста років професійні та непрофесійні дослідники намагалися її довести чи спростувати. У результаті багаторічних досліджень виявилось, що для розв'язання проблеми чотирьох фарб потрібно перевірити її справедливість для скінченного числа графів певного виду. Кількість варіантів, які потрібно перебрати, була настільки великою, що тільки за допомогою потужної ЕОМ, яка неперервно працювала більше двох місяців, у 1976 р. справедливість гіпотези чотирьох фарб було підтверджено. Однак такий "фізичний" експериментальний спосіб доведення не зовсім влаштовує багатьох професійних математиків, і вони продовжують пошуки аналітичного доведення гіпотези.

Набагато простіше можна отримати такі результати.

Теорема 4.17. Плоский граф є біхроматичним тоді й тільки тоді, коли степені всіх його граней парні.

Справедливість твердження теореми випливає з того, що в планарному графі, степені всіх граней якого парні, немає циклів непарної довжини (доведіть це самостійно). Отже, для нього виконується критерій леми 4.8.

Теорема 4.18. Для правильного розфарбування довільного планарного графа потрібно не більше шести фарб.

Доведення проведемо індукцією за кількістю n вершин графа. Для $n \leq 6$ твердження очевидне.

Припустімо, що хроматичне число всіх планарних графів із t вершинами не перевищує 6 ($t \geq 6$). Розглянемо довільний планарний граф G з $t + 1$ вершиною. Згідно з наслідком 4.13.3 в графі G існує вершина v, степінь якої не більше 5. Вилучимо вершину v з графа G. Отримаємо граф G', вершини якого за припущенням індукції можна правильно розфарбувати не більше ніж у шість кольорів. Тоді правильне розфарбування для G отримаємо з одержаного правильного розфарбування графа G', надаючи вершині v колір, відмінний від кольорів усіх суміжних із нею вершин. Оскільки таких вершин не більше п'яти, то для виконання цієї процедури достатньо шести фарб. Отже, $\chi(G) \leq 6$.

Пропонуємо самостійно переконатись у справедливості такого твердження.

Теорема 4.19. Для довільного планарного графа G виконується $\chi(G) \leq 5$.

Граф G називається ***критичним***, якщо хроматичне число підграфа G', отриманого в результаті вилучення будь-якої вершини з G, строго менше, ніж хроматичне число графа G.

Критичний граф G, для якого $k = \chi(G)$, називається ***k-критичним***.

Нарешті, зауважимо, що існують планарні графи, хроматичне число яких дорівнює 4. Найпростішим таким графом є K_4. Отже, гіпотезу чотирьох фарб не можна "вдосконалити", перетворивши в "гіпотезу трьох фарб".

Різноманітні точні та наближені алгоритми відшукання правильних розфарбувань графів можна знайти в монографії [16].

4.9. Задачі і вправи

1. Визначити хроматичне число:

(а) повного графа K_n;

(б) повного двочасткового графа $K_{n,m}$;

(в) довільного двочасткового графа;

(г) простого циклу довжини $2k$;

(д) простого циклу довжини $2k + 1$, $k \in N$;

(е) дерева.

2. Чому дорівнює хроматичне число повного графа K_n, з якого вилучено одне ребро?

3. Довести, що граф G біхроматичний тоді й тільки тоді, коли він не містить циклів непарної довжини.

4. Знайти графи, які мають різні хроматичні числа і в яких:

(а) кількість вершин степеня k однакова для всіх $k \geq 0$;

(б) кількість простих циклів довжини l однакова для всіх l;

(в) виконуються обидві умови з пунктів (а) і (б).

5. Довести, що для довільного планарного графа G виконується нерівність $\chi(G) \leq 5$.

6. Довести, що будь-який повний граф є критичним.

7. Довести, що довільний критичний граф є зв'язним.

8. Знайти всі 2-критичні та 3-критичні графи.

4.12. Обходи графів

Початок теорії графів як розділу математики пов'язують з так званою задачею про кенігсберзькі мости. Сім мостів міста Кенігсберга було розташовано на річці Прегель так, як зображено на рис. 4.11, *а*.

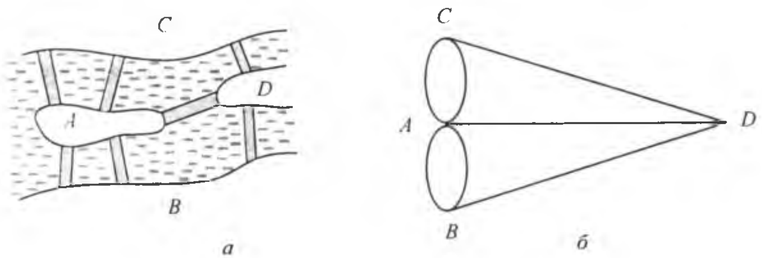


Рис. 4.11

Задача полягає в тому, чи можна, починаючи з будь-якої точки (A, B, C або D), здійснити прогулянку (обхід) через усі мости так, щоб пройти кожен міст тільки один раз і повернутися у вихідну точку.

Оскільки істотними є тільки переходи через мости, то план міста можна зобразити у вигляді графа G (з так званими *кратними* ребрами), вершинами якого є береги й острови (точки A, B, C і D), а ребрами — мости (рис. 4.11, *б*). Тоді задачу про кенігсберзькі мости мовою теорії графів можна сформулювати так: чи існує у графі G цикл, який містить усі ребра цього графа? Інше відоме формулювання цієї проблеми таке: чи можна накреслити фігуру, що зображає граф G, не відриваючи олівця від паперу та не повторюючи ліній двічі, почавши й закінчивши цю процедуру в одній із вершин фігури? Вперше відповідь на це питання дав Л. Ейлер у 1736 р. Його роботу, що містить цей розв'язок, вважають початком теорії графів.

Цикл, що містить усі ребра графа, називається ***ейлеровим циклом***. Зв'язний граф, який має ейлерів цикл, називається ***ейлеровим***.

***Теорема 4.20* (*Ейлера*).** Зв'язний граф G є ейлеровим тоді й тільки тоді, коли степені всіх його вершин парні.

Доведення. Необхідність. Нехай G — ейлерів граф. Ейлерів цикл цього графа, проходячи через кожну вершину, заходить у неї по одному ребру, а виходить — по іншому. Це означає, що кожна вершина інцидентна парній кількості ребер ейлерового циклу, а оскільки цей цикл містить усі ребра графа, то звідси випливає парність степенів усіх вершин графа.

Достатність. Припустимо, що степені всіх вершин графа $G = (V, E)$ парні. Візьмемо якусь вершину $v_1 \in V$ і розпочнемо з неї обхід графа, кожен раз обираючи ребро, яке раніше не було використано. Оскільки степінь кожної вершини парний і ненульовий, то цей шлях може закінчитися тільки у вершині v_1, утворивши таким чином

цикл Z_1. Якщо в результаті описаного процесу використано всі ребра графа G, то шуканий ейлерів цикл побудовано. Якщо ж цикл Z_1 містить не всі ребра графа G, то вилучимо з G всі ребра, які входять у Z_1. Одержимо граф G_1 — підграф графа G, усі вершини якого також матимуть парні степені (це випливає з того, що й G, і Z_1 мають вершини тільки парних степенів). Крім того, внаслідок зв'язності графа G цикл Z_1 і граф G_1 мають принаймні одну спільну вершину v_2. Відтак, починаючи з вершини v_2, побудуємо цикл Z_2 у графі G_1. Позначимо через Z_1' частину циклу Z_1 від v_1 до v_2, а через Z_1'' — частину циклу Z_1 від v_2 до v_1. Отримаємо новий цикл Z_1', Z_2, Z_1'', що веде з v_1 у v_1. Якщо цей цикл ейлерів, процес завершено. В іншому разі продовжимо аналогічні побудови ще раз і т. д. Цей процес завершиться побудовою шуканого ейлерового циклу.

Оскільки для графа G на рис. 4.11, *б* умови теореми Ейлера не виконуються, то в задачі про кенігсберзькі мости відповідь негативна.

Якщо G — ейлерів граф, то будь-який його ейлерів цикл не єдиний і може відрізнятися від інших ейлерових циклів цього графа початковою вершиною чи порядком проходження вершин (а можливо, і тим, і другим).

Для знаходження якогось ейлерового циклу в ейлеровому графі G можна застосувати так званий *алгоритм Фльорі*. Фіксуємо довільну початкову вершину циклу. На кожному кроці процедури до шуканого циклу обираємо (доки це можливо) те ребро, після вилучення якого граф не розіб'ється на дві нетривіальні зв'язні компоненти. Кожне обране ребро вилучаємо з G. Процедуру завершуємо, коли всі ребра буде вичерпано. Неважко об'рунтувати, що сформульований алгоритм будує ейлерів цикл графа G [21].

Існує ще один різновид обходу графа, який має різноманітні практичні застосування і називається *гамільтоновим циклом*. Простий цикл, який проходить через усі вершини графа, називається ***гамільтоновим циклом***. Граф називається ***гамільтоновим***, якщо він має гамільтонів цикл.

Незважаючи на певну подібність означень ейлерових і гамільтонових графів, на жаль, для розпізнавання гамільтоновості графів на сьогодні не існує таких простих і вичерпних критеріїв і алгоритмів, як для ейлерових графів. Є кілька теорем, що формулюють достатні умови існування гамільтонового циклу в заданому графі. Доведення таких теорем зазвичай містять у собі й алгоритми побудови відповідних гамільтонових циклів.

Зв'язок між ейлеровими та гамільтоновими циклами заданого графа G та відповідними циклами так званого ***реберного графа*** $L(G)$ можна встановити за допомогою нижченаведених тверджень.

Множина V' вершин графа $L(G)$ рівнопотужна множині ребер E графа G, тобто існує бієкція φ: $V' \to E$. Вершини v, $w \in V'$ з'єднано ребром у графі $L(G)$ тоді й лише тоді, коли відповідні їм ребра $\varphi(v)$ та $\varphi(w)$ суміжні у графі G.

Теорема 4.21. Якщо граф G має ейлерів цикл, то граф $L(G)$ має як ейлерів, так і гамільтонів цикли.

Теорема 4.22. Якщо граф G має гамільтонів цикл, то граф $L(G)$ також має гамільтонів цикл.

Незамкнений ланцюг, що містить усі ребра графа, називають ***ейлеровим***, а простий незамкнений ланцюг, що містить усі вершини графа, — ***гамільтоновим ланцюгом***.

4.10. Задачі і вправи

1. Довести, що зв'язний граф G є ейлеровим тоді й тільки тоді, коли G можна подати у вигляді об'єднання його підграфів, які є простими циклами.

2. Довести, що кожна вершина ейлерового графа належить якомусь циклу цього графа.

3. Визначити, які з повних графів K_n є ейлеровими.

4. Для яких значень n і m повний двочастковий граф $K_{n,m}$ є ейлеровим?

5. Довести, що зв'язний граф має ейлерів ланцюг тоді й тільки тоді, коли він має тільки дві вершини з непарними степенями.

6. Довести, що для довільного зв'язного графа існує циклічний маршрут, який починається з будь-якої вершини та містить усі ребра графа, причому кожне з них двічі.

7. Довести, що не для всіх зв'язних графів існує циклічний маршрут, який містить кожне ребро графа тричі. Сформулювати умови існування такого маршруту.

8. Довести, що в повному графі K_n існує гамільтонів цикл для довільного $n \geq 3$.

9. Довести, що граф G, який має дві несуміжні вершини степеня 3, а всі інші вершини степеня не більше 2, не має гамільтонового циклу.

10. Навести приклади ейлерового графа, який не є гамільтоновим, а також гамільтонового графа, який не є ейлеровим.

4.13. Орієнтовані графи

Крім моделі, розглянутої в попередніх розділах, у теорії досліджують й інші типи графів. Наприклад, *мультиграф* — граф, у якому дозволяються *кратні ребра*, тобто будь-які дві вершини можна з'єднати кількома ребрами. *Псевдограф* — це мультиграф, який може мати *петлі*, тобто ребра, що з'єднують вершину саму з собою. *Гіперграф* — граф, у якому ребрами можуть бути не лише двоелементні, але довільні підмножини множини вершин. Нарешті, важливою для різноманітних практичних застосувань є модель, яка називається *орієнтованим графом* (або *орграфом*). Нижче подамо короткий огляд основних понять і результатів для орграфів.

Орієнтованим графом (або ***орграфом***) G називається пара множин (V, E), де $E \subseteq V \times V$. Елементи множини V називають ***вершинами***, а елементи множини E — ***дугами орграфа*** $G = (V, E)$. Отже, дуга — це впорядкована пара вершин; V називається ***множиною вершин***, E — ***множиною дуг*** орграфа G.

Якщо $e = (v, w)$ — дуга, то вершина v називається ***початком***, а вершина w — ***кінцем*** дуги e. Кажуть, що дуга e веде з вершини v у вершину w, або виходить із v й заходить у w. Дугу e і вершини v й w називають ***інцидентними*** між собою, а вершини v і w — ***суміжними***.

Дуга (v, v), у якій початок і кінець збігаються, називається ***петлею***. Надалі розглядатимемо тільки орграфи без петель.

Як і звичайний граф, орграф $G = (V, E)$ можна задавати переліком елементів скінченних множин V і E, діаграмою чи за допомогою матриць.

Діаграма орграфа відрізняється від діаграми звичайного графа тим, що дуги орграфа зображають напрямленими лініями (відрізками чи кривими), що йдуть від початку до кінця дуги. Напрямок лінії позначають стрілкою.

Поставимо у відповідність усім вершинам орграфа $G = (V, E)$ натуральні числа від 1 до n; дістанемо множину вершин V у вигляді $\{v_1, v_2, \ldots, v_n\}$. ***Матрицею суміжності*** A орграфа G називається квадратна матриця порядку n, в якій елемент i-го рядка та j-го стовпчика a_{ij} дорівнює 1, якщо $(v_i, v_j) \in E$, і дорівнює 0 — в іншому разі.

Занумеруємо всі вершини орграфа $G = (V, E)$ числами від 1 до n, а дуги — числами від 1 до m. ***Матрицею інцидентності*** B орграфа G називається $n \times m$-матриця, у якій елемент i-го рядка та j-го стовпчика b_{ij} дорівнює 1, якщо вершина v_i є початком дуги e_j; b_{ij} дорівнює -1, якщо вершина v_i є кінцем дуги e_j; і b_{ij} дорівнює 0, якщо вершина v_i і дуга e_j неінцидентні.

Орграфи $G_1 = (V_1, E_1)$ і $G_2 = (V_2, E_2)$ називаються ***ізоморфними***, якщо існує таке взаємно однозначне відображення φ множини V_1 на множину V_2, що дуга $(v, w) \in E_1$ тоді й тільки тоді, коли дуга $(\varphi(v), \varphi(w)) \in E_2$.

Теорема 4.1 справджується й для орграфів.

Півстепенем виходу вершини v (позначається $\delta^+(v)$) орграфа G називається кількість дуг орграфа G, початком яких є вершина v. ***Півстепенем заходу вершини*** v (позначається $\delta^-(v)$) орграфа G називається кількість дуг орграфа G, кінцем яких є вершина v.

Аналогічно теоремі 4.3 можна довести таку теорему.

Теорема 4.23. Для будь-якого орграфа $G = (V, E)$ виконуються рівності $\sum_{v \in V} \delta^+(v) = \sum_{v \in V} \delta^-(v) = |E|$.

Чимало властивостей і тверджень стосовно звичайних графів можна без змін сформулювати й для орграфів. Зокрема, це стосується цілих розділів (таких, наприклад, як планарність або розфарбування графів), у яких властивість орієнтації ребер неістотна. Певні особливості в означеннях, постановках задач і методах їх розв'язання виникають при дослідженні проблем, пов'язаних із маршрутами, зв'язністю, обходами графів тощо.

Маршрутом, або ***шляхом***, в орграфі $G = (V, E)$ називається така послідовність його вершин і дуг

$$v_1, e_1, v_2, e_2, \ldots, e_k, v_{k+1}, \tag{4.5}$$

що $e_i = (v_i, v_{i+1})$, $i = 1, 2, \ldots, k$. Кажуть, що цей маршрут ***веде*** з вершини v_1 у вершину v_{k+1}. Число k дуг у маршруті (4.5) називається його ***довжиною***.

Маршрут, в якому всі дуги попарно різні, називається ***ланцюгом***, а маршрут, в якому всі вершини попарно різні, — ***простим ланцюгом***. Маршрут (4.5) називається ***замкненим*** (або ***циклічним***), якщо $v_1 = v_{k+1}$. Замкнений ланцюг називається ***циклом***, а замкнений простий ланцюг — ***простим циклом***, або ***контуром***.

Лема 4.2 виконується й для орграфів.

Орграф називається ***ациклічним*** (або ***безконтурним***), якщо він не має жодного циклу.

Якщо існує маршрут, який веде з вершини v у вершину w, то кажуть, що вершина w ***досяжна*** з вершини v. Тоді ***відстанню*** $d(v, w)$ від вершини v до вершини w називається довжина найкоротшого маршруту, що веде з v у w. Відстань між вершиною v і вершиною w, яка є недосяжною з v, позначають символом ∞.

Лема 4.9. 1). Відношення досяжності на множині вершин орграфа транзитивне.

2). Якщо в орграфі G вершина w є досяжною з вершини v, а вершина u є досяжною з вершини w, то $d(v, w) + d(w, u) \geq d(v, u)$.

Вершина v орграфа G називається ***джерелом***, якщо з неї досяжна будь-яка інша вершина орграфа G. Вершина w називається ***стоком***, якщо вона досяжна з будь-якої іншої вершини орграфа G. Вершина v орграфа G називається ***тупиковою***, якщо жодна з вершин орграфа G не досяжна з v. Вершина v орграфа G називається ***недосяжною***, якщо вона не досяжна з жодної вершини орграфа G.

Повним орграфом (або ***турніром***) називається орграф, у якому будь-які дві вершини інцидентні одній і тільки одній його дузі.

Для повних орграфів справедливі такі твердження.

Теорема 4.24. Для довільного повного орграфа $G = (V, E)$ з n вершинами виконуються такі рівності:

(a) $\sum_{i=1}^{n} \delta^+(v_i) = \sum_{i=1}^{n} (n-1-\delta^+(v_i)) = |E| = n(n-1)/2;$

(1) $\sum_{i=1}^{n} \left(\delta^+(v_i)\right)^2 = \sum_{i=1}^{n} (n-1-\delta^+(v_i))^2 = \sum_{i=1}^{n} (\delta^-(v_i))^2.$

Теорема 4.25. У будь-якому повному орграфі є принаймні одне джерело та принаймні один стік.

Теорема 4.26. У будь-якому повному орграфі існує простий ланцюг, який проходить через усі вершини орграфа.

Послідовність (4.5) називається ***напівмаршрутом***, якщо кожна дуга e_i цієї послідовності є такою, що або $e_i = (v_i, v_{i+1})$, або $e_i = (v_{i+1}, v_i)$ (можна вважати, що, будуючи нанівмаршрут, ми ігноруємо орієнтацію дуг орграфа). Аналогічно означають ***напівланцюг***, ***напівцикл*** і ***напівконтур***.

Орграф називається ***сильно зв'язним***, якщо будь-які дві його вершини досяжні одна з одної. Орграф називається ***однобічно зв'язним***, якщо для будь-яких двох його вершин принаймні одна з них досяжна з іншої. Орграф називається ***слабко зв'язним***, якщо для будь-яких двох його вершин існує напівмаршрут, що веде з однієї вершини в іншу. Маршрут в орграфі G називають ***кістяковим***, якщо він містить усі вершини орграфа G.

Сформулюємо необхідні й достатні умови для кожного з типів зв'язності.

Теорема 4.27. Орграф є сильно зв'язним тоді й тільки тоді, коли він має замкнений кістяковий маршрут.

Теорема 4.28. Орграф є однобічно зв'язним тоді й тільки тоді, коли він має кістяковий маршрут.

Теорема 4.29. Орграф є слабко зв'язним тоді й тільки тоді, коли він має кістяковий напівмаршрут.

Неважко переформулювати теореми 4.9 і 4.10, їх наслідки й алгоритми пошуку вшир і вглиб для різних типів зв'язності орграфів. Ці теореми й алгоритми можна пристосувати також для обчислення відстаней між вершинами заданого орграфа.

Орграф, у якому є джерело й немає жодного напівконтура, називається ***кореневим деревом***. ***Вхідне дерево*** — це орграф, який має стік і не має жодного напівконтура.

Орграф називається ***функціональним***, якщо півстепінь виходу кожної його вершини дорівнює 1, й ***ін'єктивним***, якщо півстепінь заходу кожної його вершини дорівнює 1.

Ейлеровим контуром в орграфі G називається контур, що містить усі дуги орграфа G. ***Ейлеровим орграфом*** називається орграф, у якому є ейлерів контур. ***Ейлеровим ланцюгом*** називається незамкнений ланцюг, що містить усі дуги орграфа.

Нижченаведену теорему можна довести так само, як і для звичайних графів.

Теорема 4.30. Слабко зв'язний орграф $G = (V, E)$ є ейлеровим тоді й тільки тоді, коли півстепінь виходу будь-якої його вершини дорівнює її півстепеню заходу.

Контур, що містить усі вершини орграфа, називається ***гамільтоновим контуром***. Орграф, який має гамільтонів контур, називається ***гамільтоновим орграфом***. Простий незамкнений ланцюг, що містить усі вершини орграфа, називається ***гамільтоновим***.

Із теореми 4.26, зокрема, випливає, що повний орграф гамільтонів.

Орграф $G = (V, E)$ називається ***транзитивним***, якщо з $(v, w) \in E$ і $(w, u) \in E$ випливає $(v, u) \in E$.

Існує взаємно однозначна відповідність між множиною всіх безконтурних транзитивних орграфів $G = (V, E)$ з множиною вершин $V = \{v_1, v_2, \ldots, v_n\}$ і петлями в кожній вершині та множиною всіх відношень часткового порядку на V. Ця бієкція встановлюється так: орграфу

$G = (V, E)$ відповідає відношення R на V таке, що $(v_i, v_j) \in R$ тоді й тільки тоді, коли $(v_i, v_j) \in E$, $v_i, v_j \in V$.

4.11. Задачі і вправи

1. Нехай задано орграф $G = (V, E)$:

(а) $V = \{1, 2, 3, 4, 5\}$, $E = \{(1, 3), (2, 1), (2, 5), (3, 4), (4, 5), (5, 1), (5, 2)\}$;

(б) $V = \{a, b, c, d\}$, $E = \{(a, c), (a, d), (b, a), (b, d), (c, b), (d, c)\}$.

Побудувати діаграму, матриці суміжності й інцидентності для кожного із заданих орграфів.

2. Довести, що для довільного орграфа $G = (V, E)$ виконується рівність:

(а) $\sum_{v \in V} \delta^+(v) = |E|$;

(1) $\sum_{v \in V} \delta^+(v) = \sum_{v \in V} \delta^-(v)$.

3. Чи існує орграф із трьома вершинами, півстепені виходу вершин якого дорівнюють 2, 2 і 0, а відповідні півстепені заходу — 2, 1 і 1?

4. Довести, що в будь-якому повному орграфі завжди є принаймні одне джерело та принаймні один стік.

5. Довести, що в повному орграфі може бути не більше однієї недосяжної і не більше однієї тупикової вершини.

6. Довести, що для довільного повного орграфа $G = (V, E)$ з n вершинами виконуються такі рівності:

(а) $$\sum_{i=1}^{n} \delta^+(v_i) = \sum_{i=1}^{n} \left(n - 1 - \delta^+(v_i)\right) = |E| = n(n-1)/2;$$

(1) $$\sum_{i=1}^{n} \left(\delta^+(v_i)\right)^2 = \sum_{i=1}^{n} \left(n - 1 - \delta^+(v_i)\right)^2 = \sum_{i=1}^{n} \left(\delta^-(v_i)\right)^2.$$

7. Довести, що в будь-якому повному орграфі існує простий ланцюг, який проходить через усі вершини орграфа.

8. Довести, що орграф сильно зв'язний тоді й тільки тоді, коли він має замкнений кістяковий маршрут.

9. Довести, що орграф однобічно зв'язний тоді й тільки тоді, коли він має кістяковий маршрут.

10. Довести, що орграф слабко зв'язний тоді й тільки тоді, коли він має кістяковий напівмаршрут.

11. ***Нумерацією*** орграфа $G = (V, E)$ з n вершинами називатимемо взаємно однозначне відображення $f\colon V \to N_n$, яке кожній вершині $v \in V$ ставить у відповідність натуральне число $f(v)$ із множини $N_n = \{1, 2, \ldots, n\}$. Довести, що для ациклічного орграфа $G = (V, E)$ існує така нумерація f, що для будь-якої дуги $(v, w) \in E$ виконується $f(v) < f(w)$. Така нумерація називається *правильною нумерацією* або *топологічним сортуванням* вершин орграфа G.

12. Довести, що для транзитивного повного орграфа завжди існує правильна нумерація.

13. Довести, що для сильно зв'язного орграфа G з n вершинами і m дугами виконуються нерівності $n \le m \le n(n-1)$.

14. Довести, що будь-який повний орграф або сильно зв'язний, або його можна перетворити в сильно зв'язний зміною орієнтації лише однієї дуги.

15. Нехай A — матриця суміжності орграфа G. Довести, що елемент $a_{ij}^{(k)}$ i-го рядка та j-го стовпчика матриці A^k дорівнює кількості шляхів довжини k, які ведуть в орграфі G з вершини з номером i у вершину з номером j.

16. Довести, що слабко зв'язний орграф $G = (V, E)$ є ейлеровим тоді й тільки тоді, коли півстепінь виходу будь-якої його вершини дорівнює її півстепеню заходу.

17. Довести, що повний орграф сильно зв'язний тоді й тільки тоді, коли він гамільтонів.

4.14. Граф як модель. Застосування теорії графів

Останнім часом графи та пов'язані з ними методи досліджень використовують практично в усіх розділах сучасної математики, зокрема дискретної.

Граф — це математична модель найрізноманітніших об'єктів, явищ і процесів, досліджуваних і використовуваних у науці, техніці та на практиці. Коротко опишемо найвідоміші застосування теорії графів.

Наприклад, у вигляді графа можна зображувати такі об'єкти:

- електричні та транспортні мережі;
- інформаційні та комп'ютерні мережі;
- карти автомобільних, залізничних і повітряних шляхів, газо- та нафтопроводів;
- моделі кристалів;
- структури молекул хімічних речовин;
- моделі ігор;
- різні математичні об'єкти (відношення, частково впорядковані множини, решітки, автомати, ланцюги Маркова, алгоритми та програми тощо);

- лабіринти;
- плани діяльності чи плани виконання певних робіт (розклади);
- генеалогічні дерева тощо.

Можна навести такі приклади застосування теорії графів:

- пошук зв'язних компонент у комунікаційних мережах;
- пошук найкоротших, "найдешевших" і "найдорожчих" шляхів у комунікаційних мережах;
- побудова кістякового дерева, тобто досягнення зв'язності з найменшою можливою кількістю ребер;
- пошук максимальної течії для транспортної мережі, у якій означено вхідні й вихідні вершини та пропускні спроможності ребер;
- ізоморфізм графів: ідентичність структур молекул (ізометрія);
- відшукання циклів графів:
  - гамільтонів цикл: обійти всі вершини графа, побувавши в кожній з них лише один раз (*задача комівояжера*);
  - ейлерів цикл: обійти всі ребра (здійснити контроль дієздатності мережі);
- розфарбування графів: розфарбування географічних карт, укладання розкладів, розміщення ресурсів тощо;
- планарність графів: проектування друкованих електронних і електричних схем, транспортних розв'язок тощо;
- знаходження центрів графа — вершин, максимальна відстань від яких до решти вершин графа мінімальна ("столиць") тощо.

РОЗДІЛ 5

КОМБІНАТОРИКА

Комбінаторика (інша назва — *комбінаторний аналіз*) — це розділ сучасної дискретної математики, що вивчає способи вибору та розміщення певних предметів, досліджує властивості і формулює методи обчислення кількостей різноманітних конфігурацій, які можна утворити з цих предметів. Оскільки конкретний вигляд (матеріальна сутність) предметів, які обирають і розміщують, не має для комбінаторного аналізу жодного значення, то при формулюванні та розв'язуванні задач комбінаторики використовують загальні поняття й терміни теорії множин і відношень. Особливо слід наголосити, що всі множини, з якими має справу комбінаторика, скінченні. Далі скрізь у цьому розділі під словами множина чи підмножина розумітимемо скінченні множини.

Користуючись мовою теорії множин, можна сказати, що комбінаторика вивчає різноманітні властивості множин, які можна утворити з підмножин певної скінченної множини. При цьому кожен раз задають певні правила, за якими формуються ці множини і підмножини.

Кількість елементів (потужність) основної скінченної множини називають ***розмірністю*** комбінаторної задачі.

Комбінаторні задачі мають давню історію. Однак тривалий час комбінаторика не привертала до себе уваги математиків. Пояснюється це значною мірою тим, що оскільки комбінаторні задачі формулюють для скінченних множин, то для переважної більшості таких задач існує тривіальний метод їх розв'язання — перебір. Пошук зручніших алгоритмів і методів для задач малої розмірності не викликає інтересу в дослідників, тому що виграш незначний порівняно з тривіальним алгоритмом перебору. Водночас комбінаторні задачі великої розмірності навіть за умови застосування найефективніших алгоритмів потребують такої кількості операцій (зокрема, такого обсягу обчислень), що стають практично нерозв'язними.

Ситуація кардинально змінилася з появою ЕОМ. Виникла реальна можливість розв'язувати комбінаторні задачі достатньо великої розмірності. Виявилося, що для таких задач різноманітні вдосконалення й оптимізація відповідних алгоритмів зумовлюють істотний виграш у часі та пам'яті. Це, у свою чергу, дає змогу додатково збільшувати розмірність задач, які можна розв'язувати за допомогою "хороших" алгоритмів. Розробка та дослідження загальних принципів побудови оптимальних комбінаторних алгоритмів (алгоритмів для розв'язування різноманітних комбінаторних задач) — одні з найважливіших проблем сучасної теорії і практики програмування [24].

5.1. Комбінаторні обчислення для основних теоретико-множинних операцій. Формула включення-виключення

Якщо A і B — довільні скінченні множини, то безпосередньо з означення теоретико-множинних операцій випливають такі співвідношення: $|A \setminus B| = |A| - |A \cap B|$, $|B \setminus A| = |B| - |A \cap B|$. Як і раніше, через $|M|$ позначено кількість елементів множини M.

Коли множини A і B не перетинаються, тобто $A \cap B = \varnothing$, то $|A \cup B| = |A| + |B|$. Зокрема, $|A \Delta B| = |A \setminus B| + |B \setminus A|$.

Теорема 5.1. Якщо $A_1, A_2, \ldots, A_n$ — скінченні множини, то

$$\begin{aligned}\Big|\bigcup_{i=1}^{n} A_i\Big| = {} & |A_1| + |A_2| + \ldots + |A_n| - (|A_1 \cap A_2| + |A_1 \cap A_3| + \ldots + |A_1 \cap A_n| + \\ & + |A_2 \cap A_3| + |A_2 \cap A_4| + \ldots + |A_{n-1} \cap A_n|) + \\ & + (|A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \ldots \\ & \ldots + |A_{n-2} \cap A_{n-1} \cap A_n|) - \ldots + (-1)^{n-1}|A_1 \cap A_2 \cap \ldots \cap A_n|. \end{aligned} \quad (5.1)$$

Доведення. Доведення проведемо методом математичної індукції.

Для $n = 2$ маємо $A_1 \cup A_2 = (A_1 \Delta A_2) \cup (A_1 \cap A_2)$. Оскільки множини $A_1 \Delta A_2$ і $A_1 \cap A_2$ не перетинаються, то $|A_1 \cup A_2| = |A_1 \Delta A_2| + |A_1 \cap A_2| = = |A_1 \setminus A_2| + |A_2 \setminus A_1| + |A_1 \cap A_2| = |A_1| - |A_1 \cap A_2| + |A_2| - |A_1 \cap A_2| + |A_1 \cap A_2| = = |A_1| + |A_2| - |A_1 \cap A_2|$.

Отже, формула (5.1) для $n = 2$ виконується.

Припустімо, що формула (5.1) справджується для $n = k$ $(k \geq 2)$

$$\left|\bigcup_{i=1}^{k} A_i\right| = S_1(A_1, A_2, \ldots, A_k) - S_2(A_1, A_2, \ldots, A_k) + S_3(A_1, A_2, \ldots, A_k) - \ldots +$$

$+ (-1)^{k-1} S_k(A_1, A_2, \ldots, A_k)$, де через $S_m(A_1, A_2, \ldots, A_p)$ позначено суму чисел $|A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}|$ за всіма можливими наборами індексів $i_1, i_2, \ldots, i_m$ такими, що $1 \leq i_1 < i_2 < \ldots < i_m \leq p$:

$$S_m(A_1, A_2, \ldots, A_p) = \sum_{1 \leq i_1 < i_2 < \ldots < i_m \leq p} \left|A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}\right|.$$

Тоді для $n = k + 1$ маємо

$$\left|\bigcup_{i=1}^{k+1} A_i\right| = \left|\bigcup_{i=1}^{k} A_i \cup A_{k+1}\right| = \left|\bigcup_{i=1}^{k} A_i\right| + |A_{k+1}| - \left|\left(\bigcup_{i=1}^{k} A_i\right) \cap A_{k+1}\right| = \left|\bigcup_{i=1}^{k} A_i\right| + |A_{k+1}| -$$

$$- \left|\bigcup_{i=1}^{k} (A_i \cap A_{k+1})\right|.$$

Застосувавши до першого та третього членів останнього виразу припущення індукції, одержимо

$$\left|\bigcup_{i=1}^{k+1} A_i\right| = S_1(A_1, A_2, \ldots, A_k) - S_2(A_1, A_2, \ldots, A_k) + \ldots + (-1)^{k-1} S_k(A_1, A_2, \ldots, A_k) +$$

$$+ |A_{k+1}| - [S_1(A_1 \cap A_{k+1}, A_2 \cap A_{k+1}, \ldots, A_k \cap A_{k+1}) - S_2(A_1 \cap A_{k+1}, A_2 \cap A_{k+1}, \ldots$$

$$\ldots, A_k \cap A_{k+1}) + \ldots + (-1)^{k-1} S_k(A_1 \cap A_{k+1}, A_2 \cap A_{k+1}, \ldots, A_k \cap A_{k+1})].$$

Використовуючи означення функцій $S_m(A_1, A_2, \ldots, A_p)$, неважко переконатись у справедливості таких співвідношень:

$S_1(A_1, A_2, \ldots, A_k) + |A_{k+1}| = S_1(A_1, A_2, \ldots, A_k, A_{k+1})$,

$S_2(A_1, A_2, \ldots, A_k) + S_1(A_1 \cap A_{k+1}, A_2 \cap A_{k+1}, \ldots, A_k \cap A_{k+1}) = S_2(A_1, A_2, \ldots, A_k, A_{k+1})$,

$S_3(A_1, A_2, \ldots, A_k) + S_2(A_1 \cap A_{k+1}, A_2 \cap A_{k+1}, \ldots, A_k \cap A_{k+1}) = S_3(A_1, A_2, \ldots, A_k, A_{k+1})$,

....

$S_k(A_1, A_2, \ldots, A_k) + S_{k-1}(A_1 \cap A_{k+1}, A_2 \cap A_{k+1}, \ldots, A_k \cap A_{k+1}) = S_k(A_1, A_2, \ldots, A_k, A_{k+1})$.

Отже, остаточно одержимо

$$\left|\bigcup_{i=1}^{k+1} A_i\right| = \sum_{i=1}^{k+1} (-1)^{i-1} S_i(A_1, A_2, \ldots, A_{k+1}).$$

Теорему доведено.

Наведемо дещо іншу інтерпретацію формули (5.1). Нехай задано якусь скінченну множину предметів M і n властивостей $p_1, p_2, \ldots, p_n$, які можуть мати ці предмети. Позначимо через $N(i_1, i_2, \ldots, i_k)$ кількість елементів множини M, які одночасно мають властивості $p_{i_1}, p_{i_2}, \ldots, p_{i_k}$. Тоді кількість елементів (предметів) множини M, які мають принаймні одну із зазначених властивостей, дорівнює

$$K = T_1 - T_2 + T_3 - \ldots + (-1)^{n-1} T_n, \tag{5.2}$$

45 $T_k = \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} N(i_1, i_2, \ldots, i_k), k = 1, 2, \ldots, n.$

Перехід від формули (5.1) до формули (5.2) можна здійснити за допомогою таких міркувань. Нехай A_i — це множина предметів з M, які мають властивість p_i. Тоді те, що елемент a має властивість p_i, можна записати у вигляді $a \in A_i$, $i = 1, 2, \ldots, n$. Множина елементів, які одночасно мають властивості $p_{i_1}, p_{i_2}, \ldots, p_{i_k}$, є $A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}$, а множина елементів, кожен із яких має принаймні одну з властивостей $p_{i_1}, p_{i_2}, \ldots, p_{i_k}$, дорівнює $A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k}$. Отже, $N(i_1, i_2, \ldots, i_k) = |A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}|$, і тотожність формул (5.1) і (5.2) стає зрозумілою.

Зауважимо, що кількість елементів множини M, які не мають жодної з розглядуваних властивостей дорівнює $|M| - K$.

Приклад 5.1. Зі 100 студентів факультету англійську мову знає 41 студент, французьку — 23, німецьку — 16, англійську та французьку — 15, англійську та німецьку — 7, французьку та німецьку — 6, усі три мови знають 3 студенти. Скільки студентів не знають жодної з трьох мов?

Позначимо через A, B, C відповідно множини студентів, які знають англійську, французьку та німецьку мову. Тоді кількість студентів, які знають принаймні одну мову, згідно з формулою (5.2), дорівнює $K = |A| + |B| + |C| - (|A \cap B| + |A \cap C| + |B \cap C|) + |A \cap B \cap C| = 41 + 23 + 16 - (15 + 7 + 6) + 3 = 55$.

Отже, шукана кількість дорівнює $100 - 55 = 45$.

Проводячи обчислення за формулами (5.1) або (5.2), потрібно послідовно додавати й віднімати певні кількості, тому метод обчислення за цими формулами дістав назву ***методу (принципу) включення*** та ***виключення***, а самі формули (5.1) і (5.2) називають ***формулами включення-виключення***.

Нарешті, розглянемо останню теоретико-множинну операцію — прямий добуток.

Теорема 5.2. Якщо $A_1, A_2, \ldots, A_n$ — скінченні множини, то

$$|A_1 \times A_2 \times \ldots \times A_n| = |A_1| \cdot |A_2| \cdot \ldots \cdot |A_n|. \quad (5.3)$$

Доведення. Як і в попередній теоремі, скористаємося методом математичної індукції.

Нехай $A_1 = \{a_1, a_2, \ldots, a_k\}$, тоді для $n = 2$ множину $A_1 \times A_2$ можна подати у вигляді

$$A_1 \times A_2 = \{a_1, a_2, \ldots, a_k\} \times A_2 = (\bigcup_{i=1}^{k}\{a_i\}) \times A_2 = \bigcup_{i=1}^{k}(\{a_i\} \times A_2).$$

Множини останнього об'єднання не перетинаються, тому

$$|A_1 \times A_2| = \sum_{i=1}^{k}|\{a_i\} \times A_2| = k|A_2| = |A_1| \cdot |A_2|.$$

Припустімо, що співвідношення (5.3) виконується для $n = k$ ($k \geq 2$). Тоді

$$|A_1 \times A_2 \times \cdots \times A_k \times A_{k+1}| = |A_1 \times A_2 \times \cdots \times A_k| \cdot |A_{k+1}| =$$
$$= |A_1| \cdot |A_2| \cdot \cdots \cdot |A_k| \cdot |A_{k+1}|.$$

Теорему доведено.

Формулу (5.3) часто називають ***основним правилом комбінаторики*** (або ***правилом множення***).

Нехай потрібно виконати одну за одною n дій. Якщо першу дію можна виконати k_1 способами, другу дію — k_2 способами і т. д., а n-ту дію — k_n способами, то всі n дій разом можна виконати $k_1 k_2 \cdots k_n$ способами.

Приклад 5.2. 1. Кількість слів довжини m в алфавіті A ($|A| = n$) дорівнює $|A^m| = |A|^m = n^m$. Цей результат випливає також із того, що побудову одного зі слів довжини m можна розкласти на m кроків (або дій): перший крок — вибір першої літери слова, другий — вибір другої літери слова і т. д., m-й крок — вибір останньої літери. Кожну з цих дій можна виконати n способами.

2. З'ясуємо, скількома способами можна розподілити k різних предметів серед n осіб.

Нехай A — множина осіб, серед яких розподіляють предмети. Кожному варіанту розподілу поставимо у відповідність кортеж $(a_{i_1}, a_{i_2}, \ldots, a_{i_k})$, де a_{i_j} — особа, яка одержала j-й предмет. Установлена відповідність взаємно однозначна, отже множина варіантів розподілу рівнопотужна множині всіх кортежів довжини k, утворених

з елементів множини A. Тому шукане число дорівнює $|A^k| = |A|^k = n^k$.

3. Нехай A і B — скінченні множини, $|A| = n$, $|B| = m$. Обчислимо кількість усіх можливих відображень множини A у множину B. Кожне відображення $\varphi: A \to B$ можна повністю задати його таблицею (табл. 5.1), де $a_1, a_2, \ldots, a_n$ — елементи множини A, а $\varphi(a_1), \varphi(a_2), \ldots$ $\ldots, \varphi(a_n)$ — відповідні образи.

Таблиця 5.1

a_1	a_2	...	a_{n-1}	a_n
$\varphi(a_1)$	$\varphi(a_2)$	...	$\varphi(a_{n-1})$	$\varphi(a_n)$

Кожен з образів можна обрати m способами. Отже, існує m^n способів утворення рядка значень для відображення φ в табл. 5.1, тому кількість відображень типу $A \to B$ дорівнює $|B^A| = m^n = |B|^{|A|}$.

5.1. Задачі і вправи

1. Нехай M — скінченна множина і $A, B \subseteq M$. Розмістіть у порядку неспадання такі величини:

(а) $|B|, |A \cup B|, |\varnothing|, |A \cap B|, |M|$;

(б) $|A \setminus B|, |A| + |B|, |A \Delta B|, |\varnothing|, |A \cup B|$.

2. Довести такі нерівності:

(а) $|A \cup B| \leq |A| + |B|$;
(б) $|A \setminus B| \leq |A|$;
(в) $|A \Delta B| \leq |A| + |B|$;
(г) $|A \setminus B| \geq |A| - |B|$;
(д) $|A \Delta B| \leq |A \cup B|$;
(е) $|A| < |\beta(A)|$;
(є) $|A \cap B| \leq |A|$;
(ж) $|A \cap B| \leq |A \cup B|$.

3. Нехай A — скінченна множина, a — елемент, а B — підмножина множини A. Яких підмножин множини A більше:

(а) тих, що містять елемент a, чи тих, що не містять елемента a;

(б) тих, що містять множину B, чи тих, що не перетинаються з множиною B;

(в) тих, що включають множину B, чи тих, що не включають множину B?

4. У групі 27 студентів. Із них 16 відвідують семінар А, 12 — семінар Б, а 7 студентів не відвідують жодного семінару:

(а) скільки студентів відвідують семінари А і Б;

(б) скільки студентів відвідують лише семінар А?

5. Обстеження читацьких смаків студентів показало, що 60 % студентів читають журнал А, 50 % — журнал Б, 50 % — журнал В, 30 % — журнали А і Б, 20 % — журнали Б і В, 40 % — журнали А і В, 10 % — журнали А, Б і В. Скільки відсотків студентів:

(а) читають принаймні один журнал;

(б) не читають жодного з журналів;

(в) читають точно два журнали;

(г) читають не менше двох журналів?

6. Знайти кількість і суму чотиризначних натуральних чисел, що не діляться на жодне з таких чисел:

(а) 3, 5, 7; (б) 6, 10, 15.

7. Знайти кількість простих чисел, що не перевищують 200.

8. Під час екзаменаційної сесії з чотирьох іспитів не менше 70 % студентів склали іспит із дискретної математики, не менше 75 % — з математичного аналізу, не менше 80 % — з алгебри та не менше 85 % — з програмування. Якою є мінімальна кількість студентів, що склали одночасно усі чотири іспити?

9. Номер автомашини складається з трьох букв українського алфавіту (що містить 33 букви) і чотирьох цифр. Скільки можна скласти різних номерів автомашин?

10. Скількома способами можна розмістити на шахівниці вісім тур так, щоб вони не били одна одну?

11. Скількома способами можна розмістити на шахівниці розміром $m \times n$ дві тури різних кольорів так, щоб вони не били одна одну?

12. Скільки існує n-значних десяткових чисел:

(а) які починаються з двох однакових цифр;

(б) у яких сусідні цифри різні;

(в) усі цифри яких непарні;

(г) у запису яких є принаймні одна парна цифра;

(д) у запису яких немає цифри 9;

(е) у запису яких обов'язково є цифра 5.

13. Скількома способами у множині A з n елементів можна вибрати дві підмножини, що не перетинаються?

14. Нехай $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ — розклад числа n на прості множники $p_1, p_2, \dots, p_r$. Знайти кількість усіх натуральних дільників числа n.

15. У скількох точках перетинаються діагоналі опуклого n-кутника, якщо жодні три з них не перетинаються в одній точці?

16. На одній прямій дано n точок, а на другій, паралельній першій, — m точок. Скільки існує трикутників, вершинами яких є ці точки?

17. Якщо повернути аркуш паперу на 180°, то цифри 0, 1 і 8 не змінюються, цифри 6 і 9 перетворюються одна в одну, а всі інші цифри втрачають

смисл. Скільки існує n-значних чисел, які при повертанні аркуша паперу на 180°:

(а) не втрачають смислу;

(б) не змінюються;

(в) не втрачають смислу й залишаються n-значними?

18. Нехай множина A містить n елементів, а множина B — m елементів. Визначити кількість:

(а) відповідностей;

(б) всюди визначених відповідностей;

(в) функціональних відповідностей;

(г) сюр'єктивних відповідностей;

(д) ін'єктивних відповідностей;

(е) бієктивних відповідностей

між множинами A і B.

19. Нехай множина M містить n елементів. Визначити кількість:

(а) відношень;

(б) рефлексивних відношень;

(в) нерефлексивних відношень;

(г) антирефлексивних відношень;

(д) неантирефлексивних відношень;

(е) нерефлексивних і неантирефлексивних відношень;

(є) симетричних відношень;

(ж) несиметричних відношень;

(з) антисиметричних відношень;

(и) рефлексивних і симетричних відношень;

(і) рефлексивних і несиметричних відношень;

(ї) рефлексивних і антисиметричних відношень;

(й) антирефлексивних і симетричних відношень;

(к) антирефлексивних і несиметричних відношень ;

(л) антирефлексивних і антисиметричних відношень

на множині M.

5.2. Сполуки, перестановки і розміщення

Позначимо через $B_k(M)$ множину всіх k-елементних підмножин даної скінченної множини M, а через $C(n, k)$ — кількість елементів множини $B_k(M)$, де $n = |M|$ і $0 \le k \le n$. Зокрема, $B_0(M)$ складається лише з одного елемента — порожньої множини $\varnothing$, $B_1(M)$ складається з n елементів — одноелементних підмножин множини M, а $B_n(M)$ містить лише один елемент — саму множину M. Отже, $C(n, 0) = C(n, n) = 1$ і $C(n, 1) = n$.

Установимо правило обчислення $C(n, k)$ для всіх інших значень k.

Щоб одержати k-елементну підмножину множини M, потрібно до $(k-1)$-елементної підмножини додати один з $n-k+1$ елементів, що не ввійшли до цієї підмножини. Оскільки кількість $(k-1)$-елементних підмножин є $C(n, k-1)$, а кожну із зазначених k-елементних підмножин можна отримати $(n-k+1)$ способом, то за допомогою описаної процедури ми дістанемо $(n-k+1)C(n, k-1)$ k-елементних підмножин. Однак не всі ці підмножини будуть різні, бо кожну з k-елементних підмножин можна так побудувати k способами: додаванням кожного з її k елементів. Тому одержане число у k разів більше, ніж шукане число $C(n, k)$. Отже,

$$k\,C(n, k) = (n-k+1)C(n, k-1). \tag{5.4}$$

Звідси послідовно знаходимо

Отже, доведено таку теорему.

$$C(n,\ k) = \frac{n-k+1}{k}C(n,\ k-1) = \frac{(n-k+1)(n-k+2)}{k(k-1)}C(n,\ k-2) = \ldots$$

$$= \frac{(n-k+1)(n-k+2)\cdots(n-1)n}{k(k-1)\cdots 2\cdot 1}C(n,\ 0) = \frac{(n-k+1)(n-k+2)\cdots(n-1)n}{k(k-1)\cdots 2\cdot 1}.$$

Теорема 5.3. Число усіх k-елементних підмножин множини M, що складається з n елементів, дорівнює

$$C(n,\ k) = \frac{(n-k+1)(n-k+2)\cdots(n-1)n}{k(k-1)\cdots 2\cdot 1} = \frac{n!}{k!(n-k)!}. \tag{5.5}$$

Довільна k-елементна підмножина n-елементної множини називається ***сполукою*** (або ***комбінацією***) ***з n елементів по k***. Формула (5.5) дає змогу обчислити кількість таких сполук.

Для кількості сполук з n по k крім уведеного позначення $C(n, k)$ використовують також позначення C_n^k, nCk, (n, k) або $\binom{n}{k}$.

Приклад 5.3. Підрахуємо, скількома способами можна заповнити картку "Лото (6 із 49)". Це число дорівнює кількості сполук із 49 елементів (чисел) по 6, тобто дорівнює

$$C(49,\ 6) = \frac{49\cdot 48\cdot 47\cdot 46\cdot 45\cdot 44}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6} = 13983816.$$

Нехай множина M містить n елементів.

Перестановкою множини M називається будь-який утворений з елементів множини M кортеж довжини n, у якому кожен елемент з M зустрічається лише один раз. Позначимо через P_n кількість усіх перестановок множини M.

Теорема 5.4. $P_n = n!$

Доведення. Будемо послідовно утворювати кортежі довжини n з елементів множини M ($|M| = n$). Є n можливостей для вибору першої координати кортежу. Після того як обрано елемент для першої координати, залишиться $n-1$ елемент, з яких можна вибрати другу координату кортежу, і т. д. За основним правилом комбінаторики всі n дій разом можна виконати $n(n-1)(n-2)\ldots 2\cdot 1 = n!$ способами. Отже, є $n!$ кортежів довжини n, утворених з елементів множини M, і $P_n = n!$

Приклад 5.4. Скільки п'ятизначних чисел можна утворити з цифр 0, 1, 2, 3, 4? Кожну цифру можна використовувати в числі тільки один раз.

За доведеною теоремою існує $P_5 = 5!$ перестановок даних цифр. Однак частина з цих перестановок матиме на першому місці цифру 0, тобто відповідатиме чотиризначним числам. Кількість чисел виду $0abcd$, де (a, b, c, d) — перестановка з цифр 1, 2, 3, 4, дорівнює $P_4 = 4!$. Отже, шукане число дорівнює $P_5 - P_4 = 5! - 4! = 96$.

Кортеж довжини k, утворений з елементів множини M ($|M| = n$), в якому елементи не повторюються, називають ***розміщенням з n по k*** ($0 \le k \le n$). Різні розміщення з n по k відрізняються або складом елементів, або їх порядком.

Кількість розміщень з n по k позначають $A(n, k)$ або A_n^k.

Теорема 5.5. $A(n, k) = k!C(n, k) = \dfrac{n!}{(n-k)!}$,

тобто $A(n, k) = n(n-1)(n-2)\ldots(n-k+1)$.

Доведення. Кількість усіх k-елементних підмножин множини M з n елементів дорівнює $C(n, k)$. Із кожної з цих підмножин можна утворити стільки кортежів, скільки існує різних перестановок її елементів. За попередньою теоремою це число дорівнює $k!$. Отже, загальна кількість кортежів довжини k, які можна побудувати з n елементів дорівнює $k!C(n, k)$.

Приклад 5.5. Визначити, скільки тризначних чисел можна утворити з цифр 0, 1, 2, 3, 4 за умови, що цифри кожного числа різні.

Існує $A(5, 3)$ кортежів довжини 3, координати яких обрано з даних цифр. Виключимо з них кортежі, перша координата яких дорівнює 0. Таких кортежів $A(4, 2)$. Отже, шукане число дорівнює $A(5, 3) - A(4, 2) = 5 \cdot 4 \cdot 3 - 4 \cdot 3 = 48$.

Нехай задано множину M, що складається з n елементів, і такі цілі числа $k_1, k_2, \ldots, k_m$, що $k_1 + k_2 + \ldots + k_m = n$ і $k_i \geq 0$, $i = 1, 2, \ldots, m$. Скільки існує розбиттів $\{A_1, A_2, \ldots, A_m\}$ множини M таких, що $|A_i| = k_i$, $i = 1, 2, \ldots, m$? Позначимо шукане число через $C_n(k_1, k_2, \ldots, k_m)$.

Теорема 5.6. $C_n(k_1, k_2, \ldots, k_m) = \dfrac{n!}{k_1!k_2!\ldots k_m!}$.

Доведення. Усі зазначені розбиття множини M на m класів можна отримати так: візьмемо довільну k_1-елементну підмножину множини M (це можна зробити $C(n, k_1)$ способами); серед $n - k_1$ елементів, що залишилися, візьмемо k_2-елементну підмножину (це можна зробити $C(n - k_1, k_2)$ способами) і т. д. Загальне число способів вибору різних множин $A_1, A_2, \ldots, A_m$ за правилом множення дорівнює

$$C(n, k_1) \cdot C(n - k_1, k_2) \cdot C(n - k_1 - k_2, k_3) \cdot \ldots \cdot C(n - k_1 - k_2 - \ldots - k_{m-1}, k_m) =$$

$$= \frac{n!}{k_1!(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \cdot \frac{(n-k_1-k_2)!}{k_3!(n-k_1-k_2-k_3)!} \cdot \ldots$$

$$\ldots \cdot \frac{(n-k_1-k_2-\ldots-k_{m-1})!}{k_m!(n-k_1-k_2-\ldots-k_m)!} = \frac{n!}{k_1!k_2! \cdots k_m!}.$$

(Нагадаємо, що $0! = 1$.)

Теорема 5.7. Число різних перестановок, які можна утворити з n елементів, серед яких є k_1 елементів першого типу, k_2 елементів другого типу і т. д., k_m елементів m-го типу, дорівнює $C_n(k_1, k_2, \ldots, k_m)$.

Доведення. Візьмемо перестановку W — довільну із зазначених перестановок — і замінимо в ній усі k_1 однакових елементів першого типу різними елементами. Тоді, виконуючи $k_1!$ перестановок нововведених елементів, одержимо $k_1!$ відповідних перестановок з n елементів. Відтак замінимо у перестановці W всі k_2 елементів другого типу різними елементами. Переставляючи ці елементи, одержимо $k_2!$ відповідних перестановок і т. д. Отже, за допомогою описаної процедури з кожної перестановки W можна утворити $k_1!k_2! \cdots k_m!$ різних перестановок довжини n. Зробивши так для всіх зазначених перестановок, дістанемо всі $n!$ можливих перестановок з n елементів. Якщо позначимо шукане число через L, то з наведених міркувань одержимо $L \cdot k_1!k_2! \ldots k_m! = n!$

Отже, $L = \dfrac{n!}{k_1!k_2!\ldots k_m!}$, тобто $L = C_n(k_1, k_2, \ldots, k_m)$.

Приклад 5.6. Скільки різних слів можна утворити, переставляючи літери слова “математика”?

Маємо 10 елементів — м, а, т, е, м, а, т, и, к, а, серед яких є два елементи “м” (першого типу), три елементи “а” (другого типу), два елементи “т” (третього типу) та по одному елементу інших типів — “е”, “и” та “к”. За доведеною теоремою з цих елементів можна утворити

$$C_{10}(2, 3, 2, 1, 1, 1) = \frac{10!}{2!3!2!1!1!1!} = 151200$$

перестановок (або слів).

5.2. Задачі і вправи

1. Скількома способами можна розподілити k екзаменаційних білетів між n студентами?

2. Скількома способами з шістьох інженерів і 14 робітників можна утворити бригаду, яка складалася б із двох інженерів і п’ятьох робітників?

3. Скількома способами можна вибрати з натуральних чисел від 1 до 100 три натуральні числа так, щоб їх сума

(а) була парною;

(б) була непарною;

(в) ділилася на три?

4. Скількома способами можна вибрати з n чоловік групу людей для роботи, якщо група має складатися не менше ніж з p чоловік?

5. Скільки є перестановок цифр 0, 1, 2, …, 9, у яких цифра 3 займає третє місце, а цифра 5 — п’яте?

6. Скількома способами можна посадити за круглий стіл n чоловіків і m жінок так, щоб жодні дві жінки не сиділи поруч?

7. Скількома способами можна розмістити n нулів і k одиниць так, щоб жодні дві одиниці не стояли поруч?

8. Нехай A і B — лінійно впорядковані множини, $|A| = n$, $|B| = m$. Відображення $f: A \to B$ називається *монотонним*, якщо з $a < b$ випливає $f(a) < f(b)$. Довести, що кількість монотонних відображень множини A в множину B дорівнює $C(n + m - 1, m - 1)$ $(m \leq n)$.

9. Скількома способами можна роздати 28 кісток доміно чотирьом гравцям так, щоб кожний одержав 7 кісток?

10. Нехай дано n символів a, m символів b і k символів c. Визначити кількість різних слів:

(а) які можна скласти з усіх цих символів;

(б) які складаються з усіх цих символів і в яких жодні два символи c не стоять поруч.

5.3. Біном Ньютона та поліномна формула

Теорема 5.8. Має місце така рівність

$$(a+b)^n = C(n,0)a^n + C(n,1)a^{n-1}b + C(n,2)a^{n-2}b^2 + \ldots + C(n,n)b^n,$$

або

$$(a+b)^n = \sum_{k=0}^{n} C(n,k)a^{n-k}b^k. \tag{5.6}$$

Доведення. Щоб переконатись у справедливості цієї рівності для будь-яких чисел a, b і n, розглянемо добуток

$$(a+b)(a+b)\ldots(a+b)\ (n\text{ множників}). \tag{5.7}$$

Розкриваючи дужки, отримаємо суму двочленів виду a^kb^{n-k}, причому двочлен a^kb^{n-k} дістанемо тоді й тільки тоді, коли з k співмножників у добутку (5.7) оберемо доданок a, а з решти $n-k$ множників — доданок b. Цей вибір можна зробити $C(n, k)$ способами. Отже, двочлен a^kb^{n-k} входить у розклад виразу (5.7) $C(n, k)$ разів, що й доводить справедливість формули (5.6).

Формулу (5.6) називають ***біномом Ньютона***, відповідну теорему — ***біномною теоремою***, а числа $C(n, k)$ — ***біномними коефіцієнтами***.

Наведемо деякі цікаві та важливі співвідношення для біномних коефіцієнтів, які називають ***біномними тотожностями***:

1) $C(n, k) = C(n, n-k)$;

2) $C(n+1, k) = C(n, k) + C(n, k-1)$;

3) $C(n, 0) + C(n, 1) + C(n, 2) + \ldots + C(n, n-1) + C(n, n) = 2^n$;

4) $C(n, 0) - C(n, 1) + C(n, 2) - \ldots + (-1)^nC(n, n) = 0$; (5.8)

5) $C(n, 0) + C(n, 2) + C(n, 4) + \ldots + C(n, k) = 2^{n-1}$, де $k = 2[n/2]$;

6) $C(n, 1) + C(n, 3) + C(n, 5) + \ldots + C(n, m) = 2^{n-1}$,

де $m = 2[(n-1)/2] + 1$.

У справедливості тотожностей 1) і 2) легко переконатися, застосовуючи, наприклад, формулу (5.5).

Тотожність 3) дістанемо, якщо в біномі Ньютона (5.6) покладемо $a = b = 1$. Вона випливає також із того, що $C(n, k)$ — це число k-елементних підмножин множини з n елементів, тому сума в лівій частині тотожності 3) дорівнює кількості всіх підмножин множини з n елементів. За теоремою 1.1 це число дорівнює 2^n.

Якщо в біномі Ньютона покласти $a = 1$ і $b = -1$, то дістанемо тотожність 4). Тотожності 5) і 6) можна отримати за допомогою відповідно додавання та віднімання тотожностей 3) та 4).

Тотожність 2) дає змогу обчислювати біномні коефіцієнти для $n + 1$, виходячи з біномних коефіцієнтів для n. Цю процедуру часто виконують за допомогою трикутної таблиці, яка називається ***трикутником Паскаля***:

1	$n = 0$
1 1	$n = 1$
1 2 1	$n = 2$
1 3 3 1	$n = 3$
1 4 6 4 1	$n = 4$
1 5 10 10 5 1	$n = 5$
....	......

У n-му рядку трикутника Паскаля стоять біномні коефіцієнти розкладу $(a + b)^n$, причому кожний коефіцієнт (окрім крайніх двох, які дорівнюють 1) дорівнює сумі двох відповідних коефіцієнтів із попереднього рядка.

Теорема 5.9 (поліномна).

$$(a_1 + a_2 + \ldots + a_m)^n = \sum_{\substack{k_1, k_2, \ldots, k_m \ge 0 \\ k_1 + k_2 + \ldots + k_m = n}} C_n(k_1, k_2, \ldots, k_m)a_1^{k_1}a_2^{k_2}\ldots a_m^{k_m}. \tag{5.9}$$

Доведення. Перемножимо $(a_1 + a_2 + \ldots + a_m)$ n разів:

$$(a_1 + a_2 + \ldots + a_m)(a_1 + a_2 + \ldots + a_m)\cdots(a_1 + a_2 + \ldots + a_m). \tag{5.10}$$

Одержимо доданки виду $a_1^{k_1}a_2^{k_2}\ldots a_m^{k_m}$ такі, що $k_i \ge 0$, $i = 1, 2, \ldots, m$, і $k_1 + k_2 + \ldots + k_m = n$. Кількість членів типу $a_1^{k_1}a_2^{k_2}\ldots a_m^{k_m}$ дорівнює кількості варіантів вибору k_1 множників із (5.10) першого типу (із цих множників у даний одночлен увійде k_1 множників a_1), відтак вибору з решти $n - k_1$ множників із (5.10) k_2 множників другого типу (із цих множників у даний одночлен увійде k_2 множників a_2) і т. д.

Отже, шукана кількість дорівнює числу розбиттів n множників у добутку (5.10) на m класів, кількість елементів у яких дорівнює відповідно $k_1, k_2, \ldots, k_m$. За теоремою 5.6 це число дорівнює

$$C_n(k_1, k_2, \ldots, k_m) = \frac{n!}{k_1!k_2!\ldots k_m!}.$$

Поліномну формулу (5.9) доведено.

Для $m = 2$ тотожність (5.9) перетворюється в біном Ньютона (5.6).

На завершення зауважимо, що насправді формула (5.6) була відома математикам задовго до Ісаака Ньютона (1642–1727). Досягненням Ньютона було узагальнення цієї формули на випадок показників, які не є натуральними числами. Він довів, що коли a — додатне число і $|x| < a$, то для довільного дійсного значення p має місце рівність

$$(a+x)^p = \frac{(p)_0}{0!}a^p + \frac{(p)_1}{1!}a^{p-1}x + \frac{(p)_2}{2!}a^{p-2}x^2 + \ldots + \frac{(p)_k}{k!}a^{p-k}x^k + \ldots,$$

або

$$(a+x)^p = \sum_{k=0}^{\infty} \frac{(p)_k}{k!}a^{p-k}x^k,$$

де $(p)_k = p(p-1)(p-2)\cdots(p-k+1)$, $k = 1, 2, \ldots$, і $(p)_0 = 1$.

Неважко переконатися, що коли p — натуральне число, остання формула перетворюється у формулу (5.6).

У зв'язку з цим узагальнюється також поняття біномного коефіцієнта. Розглядають біномні коефіцієнти $C(n, k)$ для випадку, коли n — довільне дійсне число, а k — ціле невід'ємне число. Часто, щоб відрізнити узагальнені біномні коефіцієнти від кількостей сполук $C(n, k)$ (які, очевидно, не мають смислу для ненатуральних n), їх позначають $\binom{n}{k}$. Отже, вважають, що $\binom{n}{k} = \frac{(n)_k}{k!}$. Зокрема, для натуральних n виконується рівність $\binom{n}{k} = C(n, k)$.

5.3. Задачі і вправи

1. Знайти розклад бінома:

(а) $(2x-4)^4$; (б) $(a/2+2b)^5$; (в) $(1-x^3)^5$; (г) $(a+b)^7$.

2. Знайти коефіцієнти при x^3 й x^5 у многочлені:

(а) $x(2-3x)^5 + x^3(1+2x^2)^7 - x^2(5+3x^3)^4$;

(б) $(1+x)^3 + (1+x)^4 + (1+x)^5 + \ldots + (1+x)^{15}$.

3. Знайти коефіцієнт при x^m у розкладі $(1+x)^k + (1+x)^{k+1} + \ldots + (1+x)^n$. Розглянути випадки $m < k$ і $m \geq k$.

4. Обмежившись двома членами в розкладі бінома, наближено обчислити:

(а) $(0{,}997)^8$; (б) $(2{,}003)^{10}$.

5. Знайти n, коли відомо, що:

(а) у розкладі $(1+x)^n$ коефіцієнти при x^5 та x^{12} однакові;

(б) коефіцієнти п'ятого, шостого та сьомого членів розкладу бінома $(1+x)^n$ утворюють арифметичну прогресію;

(в) восьмий член розкладу $(2x+3)^n$ має найбільший коефіцієнт.

6. Розв'язати рівняння відносно натурального k:

(а) $C(k, k-3) + C(k, k-2) = 15(k-1)$;

(б) $C(k+1, k-1) + C(k, k-2) = 9k+10$;

(в) $C(k+3, k+1) = C(k+1, k-1) + C(k, k-2) + C(k+1, k)$;

(г) $C(k, 3) + C(k, 4) = 11C(k+1, 2)$.

7. Розв'язати систему рівнянь відносно натуральних n і m:

(а) $C(n, m) = C(n, m+2)$, $C(n, 2) = 153$;

(б) $C(n+1, m-1) : C(n+1, m) = 3 : 5$, $C(n+1, m) = C(n+1, m+1)$.

8. Довести тотожність:

(а) $C(n, 1) + 6C(n, 2) + 6C(n, 3) = n^3$;

(б) $1 + 7C(n, 1) + 12C(n, 2) + 6C(n, 3) = (n+1)^3$;

(в) $C(n, 1) + 14C(n, 2) + 36C(n, 3) + 24C(n, 4) = n^4$;

(г) $C(n, 0) + 2C(n, 1) + 2^2C(n, 2) + \ldots + 2^nC(n, n) = 3^n$;

(д) $C(n, 0) + C(n, 1)/2 + \ldots + C(n, n)/n = (2^{n+1}-1)/(n+1)$;

(е) $C(n, 1) + 2C(n, 2) + \ldots + nC(n, n) = n2^{n-1}$;

(є) $C(n, 1) - 2C(n, 2) + \ldots + (-1)^{n-1}nC(n, n) = 0$;

(ж) $C(n-2, k-2) + 2C(n-3, k-2) + 3C(n-4, k-2) + \ldots + (n-k+1) \times C(k-2, k-2) = C(n, k)$;

(з) $C(n, 1) + 3C(n, 3) + 5C(n, 5) + \ldots + (2m+1)C(n, 2m+1) = n2^{n-2}$, $m = [(n-1)/2]$;

(и) $C(n, 2) + 2C(n, 4) + 3C(n, 6) + \ldots + kC(n, 2k) = n2^{n-3}$, $k = [n/2]$;

(і) $1^2C(n, 1) + 2^2C(n, 2) + 3^2C(n, 3) + \ldots + n^2C(n, n) = n(n+1)2^{n-2}$.

9. Нехай n і k — натуральні числа й $n \geq k$. Довести, що найбільший спільний дільник чисел $C(n, k)$, $C(n+1, k)$, …, $C(n+k, k)$ дорівнює 1.

10. Довести, що коли n — просте число, то всі числа n-го рядка трикутника Паскаля, окрім крайніх одиниць, діляться на n.

11. Користуючись поліномною теоремою, обчислити $(x + y + z)^3$.

12. Знайти коефіцієнти при x^k в розкладі полінома:

(а) $(1 + x^2 - x^3)^9$, $k = 8$;

(б) $(1 + x^2 + x^3)^7$, $k = 11$;

(в) $(1 + 2x^2 - 3x^4)^{10}$, $k = 8$.

13. Чому дорівнює коефіцієнт при x^n у розкладі $(1 + x + 2x^2 + \ldots + nx^n)^2$?

14. Довести, що сума всіх коефіцієнтів поліномного розкладу $(x_1 + x_2 + \ldots + x_k)^n$ дорівнює k^n.

5.4. Метод траєкторій

Існує цікава та корисна геометрична інтерпретація для біноміальних коефіцієнтів $C(n, k)$ — так званий *метод траєкторій*.

Розглянемо прямокутну розміром $m \times n$ сітку квадратів, або граф, вершинами якого є точки координатної площини з цілочисловими координатами (l, k), де $0 \le l \le m$ і $0 \le k \le n$, а ребра з'єднують вершини, що відрізняються на 1 в одній зі своїх координат (рис. 5.1).

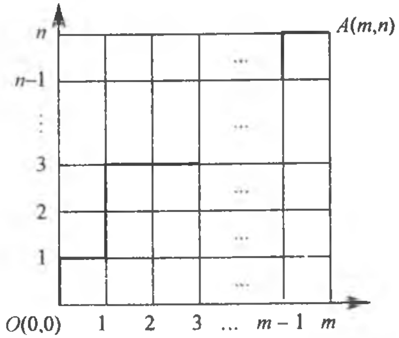


Рис. 5.1

Траєкторією назвемо будь-який шлях у цьому графі, який виходить із початкової вершини $O(0, 0)$ і такий, що на кожному кроці (з кожної вершини шляху) існує тільки два варіанти вибору для ребра шляху: вертикально вгору чи горизонтально праворуч.

На рис. 5.1 виділено одну з траєкторій, що веде з початкової вершини у вершину A з координатами (m, n). Траєкторія — це найкоротший шлях із початкової вершини у вершину A. Кожна така траєкторія складається з $m + n$ ребер, причому серед них є m горизонтальних і n вертикальних. Різні траєкторії відрізняються лише порядком чергування горизонтальних і вертикальних ребер. Тому число траєкторій дорівнює числу способів, якими з $m + n$ ребер можна вибрати n вертикальних ребер, тобто $C(n + m, n)$.

Кожен вибір n вертикальних ребер для певної траєкторії однозначно визначає варіант вибору відповідних m горизонтальних ребер. Ця відповідність взаємно однозначна. Отже, ми геометрично встановили першу з біноміальних тотожностей $C(n + m, n) = C(n + m, m)$.

Для доведення другої тотожності розглянемо рис. 5.2.

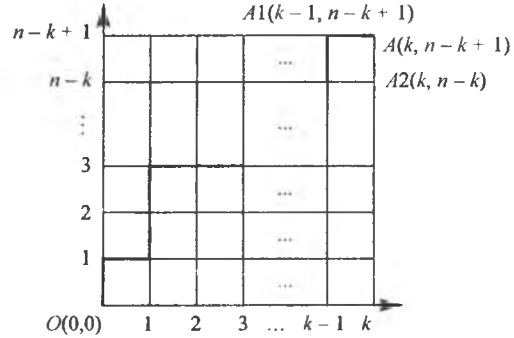


Рис. 5.2

Кількість траєкторій із вершини $O(0, 0)$ у вершину A з координатами $(k, n - k + 1)$ дорівнює $C(k + n - k + 1, k) = C(n + 1, k)$. Множину таких траєкторій можна розбити на дві підмножини: траєкторій, що проходять через точку $A1(k - 1, n - k + 1)$, і тих, що проходять через точку $A2(k, n - k)$. Кількість траєкторій першої підмножини дорівнює $C(k - 1 + n - k + 1, k - 1) = C(n, k - 1)$, а другої підмножини — $C(k + n - k, k) = C(n, k)$. Отже, ми довели, що $C(n + 1, k) = C(n, k - 1) + C(n, k)$.

Розглянемо множину всіх траєкторій довжини n на введеній вище цілочисловій координатній сітці. Кількість таких траєкторій дорівнює 2^n, бо довжина траєкторії — n, а кожне її ребро можна обрати двома способами. Кожна із зазначених траєкторій складається з k горизонтальних і $n - k$ вертикальних ребер, тому вона веде у вершину з координатами $(k, n - k)$, де $0 \le k \le n$. Множину всіх траєкторій довжини n розіб'ємо на $n + 1$ групу: траєкторії з однієї групи ведуть із початкової вершини у вершину $(k, n-k)$, $k = 0, 1, 2, \ldots, n$. Кількість траєкторій у групі з номером k дорівнює $C(k + n - k, k) = C(n, k)$, $k = 0, 1, 2, \ldots, n$. Отже, ми довели, що

$$C(n, 0) + C(n, 1) + C(n, 2) + \ldots + C(n, n-1) + C(n, n) = 2^n.$$

Користуючись методом траєкторій, доведемо таку біномну тотожність:

$$C(n, k) = C(n-1, k-1) + C(n-2, k-1) + C(n-3, k-1) + \ldots$$
$$\ldots + C(k, k-1) + C(k-1, k-1). \qquad (5.11)$$

Ліва частина — це число траєкторій, що ведуть із початкової вершини $O(0, 0)$ у цільову вершину $A(n-k, k)$. Множину цих траєкторій розіб'ємо на $n - k + 1$ груп T_m, $m = 0, 1, \ldots, n-k$. Траєкторії з групи T_m ведуть із початкової вершини в цільову та проходять через ребро, що з'єднує вершини $A_m(m, k-1)$ і $B_m(m, k)$, $m = 0, 1, \ldots, n-k$. Їх число дорівнює кількості траєкторій, що ведуть із початкової вершини в A_m, бо існує єдина траєкторія, яка веде з A_m у цільову вершину A і проходить через ребро A_mB_m. Отже, $|T_m| = C(m + k - 1, k - 1)$. Тотожність доведено.

5.4. Задачі і вправи

1. За допомогою методу траєкторій (найкоротших шляхів) довести таке співвідношення:

(а) $C(n, k) = C(n-2, k) + 2C(n-2, k-1) + C(n-2, k-2)$;

(б) $(C(n, 0))^2 + (C(n, 1))^2 + (C(n, 2))^2 + \ldots + (C(n, n))^2 = C(2n, n)$;

(в) $C(k-2, k-2)C(n-k+2, 2) + C(k-1, k-2)C(n-k+1, 2) + \ldots$ $\ldots + C(n-2, k-2)C(2, 2) = C(n, k)$;

(г) $C(n, 0)C(m, k) + C(n, 1)C(m, k-1) + \ldots + C(n, k)C(m, 0) = C(n+m, k)$, $k \le n$, $k \le m$;

(д) $C(k, k) + C(k+1, k) + C(k+2, k) + \ldots + C(n, k) = C(n+1, k+1)$, де $n \ge k$;

(е) $C(n, n) + C(n+1, n) + C(n+2, n) + \ldots + C(n+m-1, n) = C(n+m, n+1)$;

(є) $C(n-1, 0) + C(n, 1) + C(n+1, 2) + \ldots + C(n+m-1, m) = C(n+m, m)$;

(ж) $C(n, m)C(k, 0) + C(n-1, m-1)C(k+1, 1) + \ldots + C(n-m, 0) \times$ $\times C(k+m, m) = C(n+k+1, m)$.

5.5. Урнова модель

Розглянемо важливу й популярну в комбінаториці модель, яку називають *урновою моделлю*.

Нехай є k урн і n однакових предметів (наприклад, куль). Підрахуємо, скількома способами можна розподілити ці предмети по урнах. Для цього розмістимо урни послідовно та занумеруємо їх від 1 до k. Тоді кожному розподілу предметів в урнах можна поставити у відповідність кортеж $(m_1, m_2, \ldots, m_k)$ довжини k, i-та координата якого дорівнює кількості предметів, що потрапили в урну з номером i. Ця відповідність, безумовно, взаємно однозначна. Тому шукана кількість розподілів збігається з кількістю таких кортежів. Для підрахунку останньої виконаємо перетворення цих кортежів (теж взаємно однозначне). Кортеж $(m_1, m_2, \ldots, m_k)$ замінимо на кортеж із нулів і одиниць за таким правилом. Спочатку запишемо послідовність, що складається з m_1 одиниць, за нею запишемо нуль, далі — послідовність із m_2 одиниць, знову нуль і т. д. Закінчується кортеж послідовністю з m_k одиниць. Кожен отриманий у такий спосіб кортеж складається з n одиниць (оскільки $m_1 + m_2 + \ldots + m_k = n$) і $k - 1$ нулів, і множина цих кортежів взаємно однозначно відповідає множині розподілів n предметів по k урнах.

Наприклад, для $n = 7$ і $k = 3$ розподілу (5, 1, 1) (5 предметів у першій урні та по одному предмету у другій і третій) відповідає кортеж (1, 1, 1, 1, 1, 0, 1, 0, 1), а розподілу (0, 4, 3) — кортеж (0, 1, 1, 1, 1, 0, 1, 1, 1). У свою чергу, кортеж (1, 1, 1, 0, 1, 0, 1, 1, 1) — це запис розподілу, за яким у першій урні міститься 3 предмети, у другій — 1 і в третій — 3.

Кількість кортежів, що складаються з n одиниць і $k - 1$ нулів, обчислити неважко. Вона дорівнює кількості способів обрання n позицій серед $n + k - 1$ координат цих кортежів, на які буде записано одиниці (на всі інші записують нулі). Отже, шукане число дорівнює $C(n + k - 1, n)$.

До запропонованої моделі зводяться й задачі, в яких накладено певні умови на розподіли n однакових предметів по k урнах. Наприклад, якщо потрібно, щоб у результаті розподілу кожна урна містила не менше ніж t предметів ($t \ge 1$), то спочатку слід у кожну урну покласти по t предметів, відтак $n - tk$ предметів, що залишилися, розподіляти так, як і раніше. У цьому разі кількість різних розподілів дорівнює $C(n - tk + k - 1, n - tk)$ або $C(n - tk + k - 1, k - 1)$. Зокрема, якщо накладено умову, щоб при кожному розподілі жодна з урн не залишилася порожньою (тобто $t = 1$), число способів дорівнюватиме $C(n-1, n-k)$ чи $C(n-1, k-1)$.

Урнову модель можна також використати для визначення кількості невід'ємних цілих розв'язків рівняння $x_1 + x_2 + \ldots + x_k = n$, де k і n — фіксовані натуральні числа. Кожному з розв'язків $(t_1, t_2, \ldots, t_k)$

даного рівняння поставимо у відповідність певний розподіл n однакових предметів по k урнах, а саме розподіл, за яким у першу урну потрапило t_1 предметів, у другу — t_2 предметів і т. д., в останню — t_k предметів. Отже, шукана кількість розв'язків дорівнюватиме $C(n + k - 1, n)$.

Нехай задано n груп елементів, кожна з яких складається з однакових між собою елементів. ***Сполукою*** (або ***комбінацією***) ***з n по k елементів із повтореннями*** називається невпорядкований набір, що містить k елементів, причому кожний із цих елементів належить до однієї із заданих n груп.

Наприклад, якщо $n = 2$ і перша група елементів складається з літер a, а друга — з літер b, то з елементів цих двох груп можна утворити такі чотири сполуки по три елементи з повтореннями: aaa, aab, abb, bbb.

Кількість сполук з n по k елементів із повтореннями позначатимемо $\underline{C(n, k)}$. Задачу визначення цього числа можна також звести до урнової моделі, а саме до задачі розподілу k однакових предметів по n урнах. Зв'язок між сполуками та розподілами встановимо таким чином. Кожному розподілу $(m_1, m_2, \ldots, m_n)$ предметів по урнах поставимо у відповідність сполуку з повтореннями, що містить m_1 елементів першої групи, m_2 елементів другої групи і т. д., m_n елементів n-ї групи.

Отже,

$$\underline{C(n, k)} = C(n + k - 1, k) = C(n + k - 1, n - 1). \qquad (5.12)$$

Цю формулу можна отримати й інакше.

Нехай $\underline{C(n, k)}$ — число сполук з n елементів $a_1, a_2, \ldots, a_n$ по k з повтореннями. Кожна така сполука або містить a_1, або ні. Число сполук, що не містять a_1, дорівнює $\underline{C(n-1, k)}$ (це сполуки з елементів $a_2, \ldots, a_n$ по k з повтореннями). Кожна ж сполука, що містить a_1, може бути утворена приєднанням до a_1 деякої сполуки з n елементів по $k - 1$ з повтореннями (число таких сполук дорівнює $\underline{C(n, k-1)}$). Отже, матимемо

$$\underline{C(n, k)} = \underline{C(n-1, k)} + \underline{C(n, k-1)}. \qquad (5.13)$$

Послідовно застосовуючи останнє співвідношення, дістанемо

$$\underline{C(n, k)} = \underline{C(n, k-1)} + \underline{C(n-1, k)} =$$
$$= \underline{C(n, k-1)} + \left(\underline{C(n-1, k-1)} + \underline{C(n-2, k)}\right) = \qquad (5.14)$$
$$= \underline{C(n, k-1)} + \underline{C(n-1, k-1)} + \left(\underline{C(n-2, k-1)} + \underline{C(n-3, k)}\right) = \ldots$$

$$\ldots = \underline{C(n, k-1)} + \underline{C(n-1, k-1)} + \underline{C(n-2, k-1)} + \ldots$$
$$\ldots + \underline{C(2, k-1)} + \underline{C(1, k-1)}.$$

Очевидно, що $\underline{C(n, 1)} = n$ і $\underline{C(1, k)} = 1$.

Поклавши в (5.14) $k = 2$, дістанемо $\underline{C(n, 2)} = n + (n - 1) + (n - 2) + \ldots$ $\ldots + 2 + 1 = \dfrac{n(n+1)}{2} = C(n + 1, 2)$.

Для $k = 3$ матимемо:

$\underline{C(n, 3)} = C(n + 1, 2) + C(n, 2) + C(n - 1, 2) + \ldots + C(3, 2) + C(2, 2) = C(n + 2, 3)$.

У справедливості останньої рівності пропонуємо переконатися самостійно (див. задачу 5.4.1 (д)).

Повторивши подібні міркування, на $(k - 1)$-му кроці отримаємо (див. (5.11)):

$\underline{C(n, k)} = C(n + k - 2, k - 1) + C(n + k - 3, k - 1) + \ldots + C(k, k - 1) + C(k - 1, k - 1) = C(n + k - 1, k - 1)$.

Аналогічно до задач розподілу з накладеними на них додатковими умовами можна розв'язувати й задачі обчислення кількостей сполук із повтореннями, для яких мають бути виконані певні умови. Наприклад, такою умовою може бути вимога, щоб у кожну зі сполук гарантовано входило принаймні по одному елементу з t фіксованих груп та ін.

5.5. Задачі і вправи

1. Скількома способами можна розкласти 12 однакових куль по п'яти різних пакетах, якщо жоден пакет не повинен бути порожнім?

2. Скількома способами можна розкласти:

(а) 15 однакових куль по п'яти урнах так, щоб виявилося не більше двох порожніх урн;

(б) 20 однакових куль по п'яти урнах так, щоб у кожній урні виявилося не менше двох куль;

(в) 25 однакових куль по шести урнах так, щоб у кожній урні виявилося не більше п'яти куль?

3. Скільки є способів розміщення k пасажирів у n вагонах поїзда, за яких рівно m вагонів виявляться порожніми?

4. Довести, що кількість різних розв'язків рівняння $x_1 + x_2 + \ldots + x_m = n$ у натуральних числах дорівнює $C(n - 1, m - 1)$.

5. Довести, що кількість розв'язків рівняння $x_1 + x_2 + \ldots + x_m = n$ у невід'ємних цілих числах дорівнює кількості розв'язків рівняння $x_1 + x_2 + \ldots + x_m = n + m$ у натуральних числах.

6. Скільки цілих невід'ємних розв'язків має нерівність $x_1 + x_2 + \ldots + x_m \le n$?

7. Скількома способами число 11^n можна зобразити у вигляді добутку трьох співмножників (зображення, що відрізняються лише порядком співмножників, вважаються різними)?

8. Скільки розв'язків у невід'ємних цілих числах має система рівнянь $x + y = n$, $z + u + w = m$, де n, m — натуральні числа?

9. Скількома способами можна розбити n однакових куль на непорожні групи?

10. Скількома способами чотири червоні, чотири білі та чотири сині кулі можна розкласти в шість різних пакетів (деякі пакети можуть бути порожні)?

11. Скількома способами можна розмістити n білих, m червоних і k синіх куль по p урнах? Скільки є таких розміщень, що в першій урні міститься l_1 білих, l_2 червоних і l_3 синіх куль?

12. У $2n$ урнах розподілено n білих і n чорних однакових куль, причому в кожній урні є принаймні одна куля. Скількома різними способами це можна зробити?

13. Скількома способами можна розмістити $2n$ білих, $2n$ червоних і $2n$ синіх куль у двох урнах так, щоб у кожній урні було $3n$ куль?

14. Скількома способами можна вибрати k тістечок у кондитерській, де є n різних сортів тістечок?

15. Виписано всі сполуки з повтореннями з n по m елементів. Скільки разів у них повторюється кожен елемент?

5.6. Рекурентні співвідношення

У багатьох розділах математики, зокрема в комбінаториці, обчислення значень певної функції часто можна звести до обчислення значень цієї функції для менших значень параметрів.

Формулу (чи правило), що задає процедуру зведення, називають ***рекурентним***, або ***зворотним***, ***співвідношенням***.

Приклад 5.7. 1. Для довільних членів арифметичної та геометричної прогресій виконуються відповідно такі рекурентні співвідношення: $a_n = a_{n-1} + d$ та $b_n = b_{n-1}q$. Додатково для арифметичної прогресії можна отримати інше рекурентне співвідношення: $a_n = 2a_{n-1} - a_{n-2}$.

2. Популярним рекурентним співвідношенням є співвідношення, що задає так звані *числа Фібоначчі*:

$$F(n) = F(n-1) + F(n-2),\ F(0) = F(1) = 1. \qquad (5.15)$$

3. Вище ми отримали рекурентне співвідношення для чисел $C(n, k)$ (див. (5.4)):

$$C(n, k) = \frac{n-k+1}{k}\ C(n, k-1). \qquad (5.16)$$

4. Неважко довести таке рекурентне співвідношення для числа перестановок з n елементів: $P_n = P_{n-1}n$.

Розв'язком рекурентного співвідношення називається функція, що задовольняє це співвідношення, тобто така, підставивши яку в співвідношення, отримаємо тотожність.

Наприклад, розв'язками рекурентних співвідношень для арифметичної та геометричної прогресій є відповідно функції $g_1(n) = a_1 + (n-1)d$ та $g_2(n) = b_1q^{n-1}$, розв'язок рекурентного співвідношення (5.16) дає формула (5.5), а розв'язок рекурентного співвідношення для числа перестановок з n елементів — це функція $g(n) = n!$.

Загального універсального методу для розв'язування довільних рекурентних співвідношень, на жаль, немає. Однак для певного досить поширеного класу рекурентних співвідношень такий метод є. Ці співвідношення мають вигляд

$$f(n) = a_1f(n-1) + a_2f(n-2) + \ldots + a_kf(n-k), \qquad (5.17)$$

де $a_i \in R$, $i = 1, 2, \ldots, k$, і називаються ***лінійними однорідними рекурентними співвідношеннями з постійними (сталими) коефіцієнтами***. Число k називають ***порядком***, або ***глибиною***, рекурентного співвідношення (5.17). Часто замість $f(m)$ записують f_m.

Опишемо метод розв'язання таких рекурентних співвідношень.

Теорема 5.10. Якщо функції $g_1(n)$, $g_2(n)$, ..., $g_m(n)$ є розв'язками рекурентного співвідношення (5.17), то для будь-яких чисел $C_1, C_2, \ldots, C_m$ функція

$$g(n) = C_1g_1(n) + C_2g_2(n) + \ldots + C_mg_m(n)$$

також є розв'язком цього співвідношення.

Доведення. Оскільки функції $g_1(n)$, $g_2(n)$, ..., $g_m(n)$ — розв'язки співвідношення (5.17), то для них виконуються тотожності

$$g_i(n) = a_1g_i(n-1) + a_2g_i(n-2) + \ldots + a_kg_i(n-k), \quad i = 1, 2, \ldots, m.$$

Помноживши обидві частини цих тотожностей відповідно на $C_1, C_2, \ldots, C_m$ і додавши результати, дістанемо, що функція $g(n)$ також задовольняє співвідношення (5.17).

Теорема 5.11. Якщо λ_1 — корінь рівняння

$$\lambda^k = a_1\lambda^{k-1} + a_2\lambda^{k-2} + \ldots + a_{k-1}\lambda + a_k, \tag{5.18}$$

то функція $g(n) = \lambda_1^n$ є розв'язком рекурентного співвідношення (5.17).

Доведення. Справді, підставивши функцію $g(n) = \lambda_1^n$ у співвідношення (5.17), отримаємо рівність

$$\lambda_1^n = a_1\lambda_1^{n-1} + a_2\lambda_1^{n-2} + \ldots + a_{k-1}\lambda_1^{n-k+1} + a_k\lambda_1^{n-k}.$$

Вона справджується, оскільки λ_1 — корінь рівняння (5.18).

Рівняння (5.18) називають ***характеристичним рівнянням*** лінійного однорідного рекурентного співвідношення з постійними коефіцієнтами (5.17).

Корені рівняння (5.18) можуть бути простими та кратними. Розглянемо спочатку випадок, коли всі корені $\lambda_1, \lambda_2, \ldots, \lambda_k$ характеристичного рівняння прості. Тоді за теоремою 5.11 функції $\lambda_1^n, \lambda_2^n, \ldots, \lambda_k^n$ є розв'язками рекурентного співвідношення (5.17).

Теорема 5.12. Довільний розв'язок $g(n)$ рекурентного співвідношення (5.17), характеристичне рівняння якого має прості корені $\lambda_1, \lambda_2, \ldots, \lambda_k$, можна подати у вигляді

$$g(n) = C_1\lambda_1^n + C_2\lambda_2^n + \ldots + C_k\lambda_k^n. \tag{5.19}$$

Доведення. Будь-який розв'язок $g(n)$ рекурентного співвідношення (5.17) повністю й однозначно визначений, якщо задано значення $g(0), g(1), \ldots, g(k-1)$. Отже, для доведення існування формули (5.19) потрібно довести, що система рівнянь

$$\begin{cases} C_1 + C_2 + \ldots + C_k = g(0), \\ C_1\lambda_1 + C_2\lambda_2 + \ldots + C_k\lambda_k = g(1), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ C_1\lambda_1^{k-1} + C_2\lambda_2^{k-1} + \ldots + C_k\lambda_k^{k-1} = g(k-1) \end{cases} \tag{5.20}$$

має єдиний розв'язок відносно $C_1, C_2, \ldots, C_k$ для довільних $g(0), g(1), \ldots, g(k-1)$.

Визначник системи (5.20) є відомим *визначником Вандермонда*

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_k \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_1^{k-1} & \lambda_2^{k-2} & \ldots & \lambda_k^{k-1} \end{vmatrix},$$

що дорівнює добутку $\prod_{1\le i<j\le k}(\lambda_i - \lambda_j)$. Оскільки всі λ_i попарно різні, то цей визначник не дорівнює 0, тому система (5.20) крамерівська, тобто сумісна і має єдиний розв'язок. Теорему доведено.

Відтак розглянемо випадок, коли характеристичне рівняння рекурентного співвідношення (5.17) має кратні корені.

Нехай характеристичне рівняння (5.18) має s попарно різних коренів $\lambda_1, \lambda_2, \ldots, \lambda_s$, кратності яких дорівнюють відповідно $l_1, l_2, \ldots, l_s$, де $l_1 + l_2 + \ldots + l_s = k$. Можна довести, що в цьому випадку окрім функцій $\lambda_1^n, \lambda_2^n, \ldots, \lambda_s^n$ розв'язками рекурентного співвідношення (5.17) будуть також функції $n\lambda_j^n, n^2\lambda_j^n, \ldots, n^{l_j-1}\lambda_j^n$ для $j = 1, 2, \ldots, s$.

Теорема 5.13. Довільний розв'язок $g(n)$ рекурентного співвідношення (5.17), характеристичне рівняння якого має s різних коренів $\lambda_1, \lambda_2, \ldots, \lambda_s$ відповідно кратностей $l_1, l_2, \ldots, l_s$, можна подати у вигляді

$$g(n) = \sum_{j=1}^{s}(C_{j1} + C_{j2}n + C_{j3}n^2 + \ldots + C_{jl_j}n^{l_j-1})\lambda_j^n. \tag{5.21}$$

Вирази (5.19) і (5.21) називають загальними розв'язками лінійного однорідного рекурентного співвідношення з постійними коефіцієнтами. Неважко переконатися, що формула (5.19) є окремим випадком формули (5.21).

Отже, коли характеристичне рівняння (5.18) має тільки прості корені $\lambda_1, \lambda_2, \ldots, \lambda_k$, набір функцій $\lambda_1^n, \lambda_2^n, \ldots, \lambda_k^n$ утворює базис у просторі всіх розв'язків рекурентного співвідношення (5.17). Для кореня λ_j кратності l_j функціями базису будуть $\lambda_j^n, n\lambda_j^n, n^2\lambda_j^n, \ldots, n^{l_j-1}\lambda_j^n, j = 1, 2, \ldots, s$.

Приклад 5.8. Розглянемо рекурентне співвідношення для чисел Фібоначчі: $F(n) = F(n-1) + F(n-2)$. Його характеристичне рівняння має вигляд $\lambda^2 = \lambda + 1$. Коренями цього квадратного рівняння є числа $\lambda_1 = (1+\sqrt{5})/2$ і $\lambda_2 = (1-\sqrt{5})/2$. Отже, довільний розв'язок даного рекурентного співвідношення можна записати у вигляді

$$g(n) = C_1\left(\frac{1+\sqrt{5}}{2}\right)^n + C_2\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Оскільки два перші числа Фібоначчі фіксують і вважають рівними 1, то $g(0) = g(1) = 1$. Тоді для параметрів C_1 і C_2 отримаємо таку систему рівнянь:

$$\begin{cases} C_1 + C_2 = 1, \\ C_1\left(\frac{1+\sqrt{5}}{2}\right) + C_2\left(\frac{1-\sqrt{5}}{2}\right) = 1. \end{cases}$$

Її розв'язок: $C_1 = \frac{\sqrt{5}+1}{2\sqrt{5}}, C_2 = \frac{\sqrt{5}-1}{2\sqrt{5}}$.

Отже, довільне число Фібоначчі можна визначити за формулою

$$F(n) = \frac{1}{2\sqrt{5}}\left((\sqrt{5}+1)\left(\frac{1+\sqrt{5}}{2}\right)^n + (\sqrt{5}-1)\left(\frac{1-\sqrt{5}}{2}\right)^n\right),$$

або

$$F(n) = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right). \quad (5.22)$$

Нарешті, розглянемо *лінійні неоднорідні рекурентні співвідношення з постійними коефіцієнтами*, які мають вигляд

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \ldots + a_k f(n-k) + h(n), \quad (5.23)$$

де $h(n)$ — відома функція.

Теорема 5.14. Довільний (загальний) розв'язок лінійного неоднорідного рекурентного співвідношення (5.23) є сумою деякого частинного його розв'язку та загального розв'язку $g(n)$ відповідного лінійного однорідного співвідношення.

Доведення. Нехай $G(n)$ — довільний частинний розв'язок неоднорідного рекурентного співвідношення (5.23), а $g(n)$ — загальний розв'язок відповідного лінійного однорідного співвідношення (5.17). Підставивши функцію $G(n) + g(n)$ у співвідношення (5.23), дістанемо

$$G(n) + g(n) = \sum_{i=1}^{k}(a_i(G(n-i) + g(n-i)) + h(n).$$

Оскільки $G(n)$ — розв'язок неоднорідного співвідношення, то

$$G(n) = \sum_{i=1}^{k} a_i G(n-i) + h(n).$$

Отже, для функції $g(n)$ має виконуватись рівність $g(n) = \sum_{i=1}^{k} a_i g(n-i)$, а це означає, що функція $g(n)$ — розв'язок відповідного однорідного рекурентного співвідношення.

Остання теорема зводить задачу відшукання загального розв'язку неоднорідного рекурентного співвідношення з постійними коефіцієнтами (5.23) до задачі знаходження будь-якого частинного його розв'язку та відомої проблеми розв'язання однорідного рекурентного співвідношення.

Є різні методи відшукання частинного розв'язку неоднорідного співвідношення залежно від типу функції $h(n)$.

Якщо функція $h(n)$ не залежить від n, тобто є деякою дійсною константою b, то частинний розв'язок знаходимо в такий спосіб. Коли $A = a_1 + a_2 + \ldots + a_k \neq 1$, то частинним розв'язком буде $G(n) = c$, де $c = b/(1 - A)$. Якщо ж $A = 1$ і $B = 1 \cdot a_1 + 2a_2 + \ldots + ka_k \neq 0$, то існує частинний розв'язок $G(n) = dn$, де $d = b/B$.

Складнішою є задача, коли функція $h(n)$ істотно залежить від n.

Зокрема, якщо $h(n)$ — многочлен степеня t, а число 1 — корінь характеристичного рівняння кратності l, то частинний розв'язок слід шукати у вигляді $G(n) = b_1 n^l + b_2 n^{l+1} + \ldots + b_{t+1} n^{l+t}$; коефіцієнти $b_1, b_2, \ldots, b_{t+1}$ можна визначити, підставивши $G(n)$ у співвідношення (5.23) і використавши відомі початкові значення функції $f(n)$ для $n = 1, 2, \ldots, k$.

Приклад 5.9. Розв'язати рекурентне співвідношення $f_{n+3} = 3f_{n+2} - 4f_n$ з початковими умовами $f_1 = 2, f_2 = -2, f_3 = 6$.

Запишемо його характеристичне рівняння $\lambda^3 = 3\lambda^2 + 0 \cdot \lambda - 4$, тобто $\lambda^3 - 3\lambda^2 + 4 = 0$. Його коренями будуть числа $\lambda_1 = -1$ і $\lambda_{2,3} = 2$; отже, довільний розв'язок даного рекурентного співвідношення має вигляд $f_n = C_1(-1)^n + (C_2 + C_3 n) \cdot 2^n$.

Використовуючи початкові умови, дістанемо систему рівнянь

$$\begin{cases} f_1 = \ \ 2 = -C_1 + 2C_2 + \ 2C_3, \\ f_2 = -2 = \ C_1 + 4C_2 + \ 8C_3, \\ f_3 = \ \ 6 = -C_1 + 8C_2 + 24C_3. \end{cases}$$

Звідси $C_1 = 1/3$, $C_2 = -5/9$, $C_3 = -22/9$. Остаточно маємо:

$$f_n = (3(-1)^n - (5 + 22n)2^n)/9.$$

Приклад 5.10. Розв'язати рекурентне співвідношення

$f_{n+3} = -f_{n+2} + 5f_{n+1} - 3f_n + (2n^2 - n - 1)$, де $f_1 = -1, f_2 = 3, f_3 = 5$.

Характеристичне рівняння відповідного однорідного рекурентного співвідношення має вигляд $\lambda^3 = -\lambda^2 + 5\lambda - 3$, або $\lambda^3 + \lambda^2 - 5\lambda + 3 = 0$; його корені — $\lambda_{1,2} = 1$, $\lambda_3 = -3$. Звідси отримуємо загальний розв'язок цього співвідношення: $f_n = (C_1 + C_2 n) \cdot 1^n + C_3(-3)^n = C_1 + C_2 n + C_3(-3)^n$.

Частинний розв'язок шукатимемо у вигляді $G(n) = b_1 n^2 + b_2 n^3 + b_3 n^4$. Для знаходження b_1, b_2 та b_3 підставимо $G(n)$ у початкове рівняння

$$b_1(n+3)^2 + b_2(n+3)^3 + b_3(n+3)^4 = -\left(b_1(n+2)^2 + b_2(n+2)^3 + \right.$$
$$\left. + b_3(n+2)^4\right) + 5\left(b_1(n+1)^2 + b_2(n+1)^3 + b_3(n+1)^4\right) -$$
$$-3\left(b_1 n^2 + b_2 n^3 + b_3 n^4\right) + (2n^2 - n - 1).$$

Застосувавши метод невизначених коефіцієнтів, послідовно знайдемо b_1, b_2 і b_3, а згодом і C_1, C_2, C_3.

Теорема 5.15. Якщо для послідовності чисел $x_1, x_2, \ldots, x_n, \ldots$ виконується лінійне однорідне рекурентне співвідношення з постійними коефіцієнтами

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \ldots + a_k x_{n-k}, \tag{5.24}$$

то для послідовності сум $S_m = x_1 + x_2 + \ldots + x_m$ виконується лінійне однорідне рекурентне співвідношення з постійними коефіцієнтами

$$S_n = (1 + a_1)S_{n-1} + (a_2 - a_1)S_{n-2} + \ldots + (a_k - a_{k-1})S_{n-k} - a_k S_{n-k-1}.$$

Доведення. Зауважимо, що $x_i = S_i - S_{i-1}$, $i = 1, 2, \ldots$. Підставивши ці значення у співвідношення (5.24), маємо

$$S_n - S_{n-1} = a_1(S_{n-1} - S_{n-2}) + a_2(S_{n-2} - S_{n-3}) + \ldots + a_k(S_{n-k} - S_{n-k-1}).$$

З останньої рівності, виконавши нескладні перетворення, отримаємо шукане рекурентне співвідношення для послідовності S_n.

Приклад 5.11. Розглянемо послідовність кубів натуральних чисел 1, 8, 27, 64, ...

Неважко переконатися, що члени цієї послідовності задовольняють рекурентне співвідношення $x_n = 4x_{n-1} - 6x_{n-2} + 4x_{n-3} - x_{n-4}$. Тоді за доведеною теоремою для послідовності сум $S_n = \sum_{i=1}^{n} i^3$ виконується рекурентне співвідношення

$$S_n = 5S_{n-1} - 10S_{n-2} + 10S_{n-3} - 5S_{n-4} + S_{n-5}. \tag{5.25}$$

Розв'яжемо його.

Відповідне характеристичне рівняння має вигляд

$$\lambda^5 - 5\lambda^4 + 10\lambda^3 - 10\lambda^2 + 5\lambda - 1 = 0.$$

Його корінь $\lambda = 1$ має кратність 5. Отже, загальним розв'язком для співвідношення (5.25) є

$$g(n) = C_1 n^4 + C_2 n^3 + C_3 n^2 + C_4 n + C_5.$$

Використовуючи початкові умови для послідовності S_i ($S_1 = 1$, $S_2 = 9$, $S_3 = 36$, $S_4 = 100$, $S_5 = 225$), дістанемо систему лінійних рівнянь

$$\begin{cases} C_1 + C_2 + C_3 + C_4 + C_5 = 1, \\ 16C_1 + 8C_2 + 4C_3 + 2C_4 + C_5 = 9, \\ 81C_1 + 27C_2 + 9C_3 + 3C_4 + C_5 = 36, \\ 256C_1 + 64C_2 + 16C_3 + 4C_4 + C_5 = 100, \\ 625C_1 + 125C_2 + 25C_3 + 5C_4 + C_5 = 225. \end{cases}$$

Її розв'язок: $C_1 = 1/4$, $C_2 = 1/2$, $C_3 = 1/4$, $C_4 = C_5 = 0$. Отже,

$$g(n) = \frac{n^4}{4} + \frac{n^3}{2} + \frac{n^2}{4} = \frac{n^2(n^2 + 2n + 1)}{4} = \left(\frac{n(n+1)}{2}\right)^2,$$

тобто

$$\sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

5.6. Задачі і вправи

1. Знайти a_6 і a_7, якщо

(а) $a_n = n^2 + 2a_{n-1}$ для $n \ge 2$ і $a_1 = 3$;

(б) $a_n = a_{n-1}^2 - n a_{n-2}$ для $n \ge 3$ і $a_1 = a_2 = 1$;

(в) $a_n = a_{n-1}$, якщо $n = 2k$; $a_n = a_{n-1} + a_{n-2}$, якщо $n = 2k + 1$, $k \in N$ і $a_1 = 2$.

2. Розв'язати рекурентне співвідношення:

(0) $a_n = \dfrac{a_{n-1}}{1 + a_{n-1}}$, $a_1 = 1$, $n \ge 2$;

(1) $a_n = \sqrt{a_{n-1}^2 + 1}$, $a_1 = 1$, $n \ge 2$.

3. Довести, що коли послідовність $\{a_n\}$ задовольняє співвідношення $a_n = Aa_{n-1} + Ba_{n-2} + C$, то послідовність $b_n = a_n - a_{n-1}$ задовольняє співвідношення $b_n = Ab_{n-1} + Bb_{n-2}$.

4. Знайти загальний розв'язок рекурентного співвідношення:

(а) $a_{n+2} + 7a_{n+1} + 12a_n = 0$;

(б) $a_{n+2} - 4a_{n+1} + 13a_n = 0$;

(в) $a_{n+2} + 9a_n = 0$;

(г) $a_{n+3} - 9a_{n+2} + 26a_{n+1} - 24a_n = 0$;

(д) $a_{n+2} + 2a_{n+1} + a_n = 0$;

(е) $a_{n+4} + 4a_n = 0$.

5. Знайти загальний розв'язок рекурентного співвідношення із заданими початковими умовами:

(а) $a_{n+2} - 5a_{n+1} + 6a_n = 0$, $a_1 = 1$, $a_2 = -7$;
(б) $a_{n+3} - 3a_{n+2} + a_{n+1} - 3a_n = 0$, $a_1 = 3$, $a_2 = 7$, $a_3 = 27$;
(в) $a_{n+3} - 3a_{n+2} - 10a_{n+1} + 24a_n = 0$, $a_1 = 4$, $a_2 = 46$, $a_3 = 34$;
(г) $a_{n+3} - 3a_{n+1} + 2a_n = 0$, $a_1 = a$, $a_2 = b$, $a_3 = c$;
(д) $a_{n+2} - 2\cos\alpha \cdot a_{n+1} + a_n = 0$, $a_1 = \cos\alpha$, $a_2 = \cos 2\alpha$;
(е) $a_{n+2}^2 - 5a_{n+1}^2 + 4a_n^2 = 0$, $a_1 = 4$, $a_2 = 13$.

6. Розв'язати лінійне неоднорідне рекурентне співвідношення:

(а) $a_{n+1} - a_n = n$, $a_1 = 1$;
(б) $a_{n+3} - 6a_{n+2} + 11a_{n+1} - 6a_n = 6n^2 - 4n - 17$, $a_1 = 3$, $a_2 = 15$, $a_3 = 41$.

7. Використовуючи теорему 5.15 і приклад 5.11, знайти формулу для обчислення суми:

(а) $1^2 + 2^2 + \ldots + n^2$; (б) $1^4 + 2^4 + \ldots + n^4$.

8. Для послідовності чисел Фібоначчі $\{F_n\}$ довести, що:

(а) для будь-яких натуральних чисел m і n $F_{n+m} = F_{n-1}F_m + F_nF_{m+1}$;
(б) для будь-яких чисел m і $n = km$ число F_n ділиться на F_m;
(в) два сусідні числа F_n і F_{n+1} взаємно прості;
(г) $F_1 + F_3 + \ldots + F_{2n+1} = F_{2n+2}$;
(д) $1 + F_2 + F_4 + \ldots + F_{2n} = F_{2n+1}$;
(е) $F_n + F_{n+1} + F_{n+2} + F_{n+3} + F_{n+4} + F_{n+5} = 4F_{n+4}$;
(є) $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$, $n > 1$;
(ж) $C(n, 1)F_1 + C(n, 2)F_2 + \ldots + C(n, n)F_n = F_{2n}$.

5.7. Метод генератрис

Метод генератрис, або твірних функцій, є одним із теоретично й аналітично найрозвиненіших методів комбінаторного аналізу, та одним з тих, які найтісніше пов'язують комбінаторику з іншими розділами математики, зокрема з математичним аналізом. Головні ідеї цього методу вперше висловив і застосував усередині XVIII століття Леонард Ейлер у дослідженнях із теорії чисел. Згодом цей підхід успішно та плідно розвинув видатний французький вчений П'єр Сімон Лаплас (1749–1827).

Розглянемо спочатку прості приклади, які, можливо, лежали в основі перших відкриттів цієї теорії. Нехай маємо скінченний набір дійсних чисел $x_1, x_2, \ldots, x_n$ і добуток лінійних біномів:

$$\prod_{i=1}^{n}(1 + x_i t) = \sum_{m=0}^{n} a_m t^m. \qquad (5.26)$$

Неважко переконатися, що

$a_0 = 1$,
$a_1 = x_1 + x_2 + \ldots + x_n$,
$a_2 = x_1x_2 + x_1x_3 + \ldots + x_1x_n + x_2x_3 + x_2x_4 + \ldots + x_2x_n + \ldots + x_{n-1}x_n$,
$a_3 = x_1x_2x_3 + x_1x_2x_4 + \ldots + x_1x_2x_n + x_1x_3x_4 + x_1x_3x_5 + \ldots + x_{n-2}x_{n-1}x_n$,
....
$a_n = x_1x_2 \cdots x_n$.

Отже, коефіцієнт a_m — це симетрична функція, складена з усіх можливих добутків елементів x_i, узятих по m штук, тобто з добутків елементів усіх сполук з n по m множини $\{x_1, x_2, \ldots, x_n\}$. Зокрема, коли $x_i = 1$ для всіх $i = 1, 2, \ldots, n$, то коефіцієнт a_m дорівнюватиме числу $C(n, m)$ сполук з n по m ($m = 0, 1, 2, \ldots, n$) і виконуватиметься рівність:

$$(1+t)^n = \sum_{m=0}^{n} C(n,m)t^m. \qquad (5.27)$$

Вираз (5.27) взаємно однозначно пов'язує послідовності чисел (біноміальних коефіцієнтів) $\{C(n, m) \mid m = 0, 1, 2, \ldots, n\}$ із функціями виду $F(t) = (1 + t)^n$, $n = 0, 1, 2, \ldots$. Цей зв'язок виявляється дуже корисним: надаючи у формулі (5.27) різні значення змінній t, можна одержувати різні співвідношення для біномних коефіцієнтів, важливі й цікаві як для практичних застосувань, так і з теоретичної точки зору.

Наприклад, для $t = 1$ і $t = -1$ маємо відповідно співвідношення 3) і 4) з (5.8).

Розбиття функції $F(t)$ на два співмножники $(1 + t)^n = (1 + t)^k(1 + t)^{n-k}$ дає змогу отримати ще одне важливе співвідношення:

$$C(n,m) = \sum_{r=0}^{m} C(k,m-r)C(n-k,r), m = 0,1,2,\ldots,n. \qquad (5.28)$$

Окремий випадок цієї рівності дістанемо для парного $n = 2k$, якщо розіб'ємо $(1 + t)^n$ на однакові множники, тобто $m = k$. Тоді матимемо

$$C(2k,k) = \sum_{r=0}^{k}\left(C(k,r)\right)^2.$$

Продиференціювавши обидві частини співвідношення (5.27) за змінною t, одержимо рівність:

$$n(1+t)^{n-1} = \sum_{m=0}^{n} mC(n,m)t^{m-1}.$$

Підставивши в цю рівність замість t одиницю, дістанемо таке співвідношення:

$$\sum_{m=0}^{n} mC(n,m) = n \cdot 2^{n-1}. \tag{5.29}$$

Продовжуючи виконувати різні аналітичні операції над функцією $F(t)$ (наприклад, можна визначити її другу, третю і т. д. похідну, знайти її первісну тощо), отримуватимемо нові співвідношення для біномних коефіцієнтів.

Функцію $F(t) = (1 + t)^n$ називають *генератрисою* послідовності чисел $(C(n, 0), C(n, 1), C(n, 2), \ldots, C(n, n))$.

Поняття генератриси можна узагальнити в такий спосіб.

Нехай маємо деяку числову послідовність $(a_0, a_1, a_2, \ldots, a_n, \ldots)$. ***Генератрисою*** (або ***твірною функцією***) цієї послідовності називається функція

$$F(t) = \sum_{n=0}^{\infty} a_n t^n. \tag{5.30}$$

Якщо цей ряд збігається в деякій області й функцію $F(t)$ можна подати в аналітичній формі (наприклад, у вигляді формули), то за допомогою генератриси $F(t)$ вдається зручно й ефективно досліджувати різні властивості відповідної послідовності. Нас більше цікавитимуть генератриси для послідовностей, так чи інакше пов'язаних із комбінаторними задачами.

Зауважимо, що для багатьох комбінаторних задач збіжність ряду (5.30) неістотна, тому його часто називають "формальним рядом". У комбінаториці замість обчислення суми цього ряду для конкретного значення змінної t виконують певні формальні операції для таких рядів і/або визначають коефіцієнти при окремих степенях t.

Доведемо, наприклад, що генератрисою послідовності біномних коефіцієнтів $(C(n-1, 0), C(n, 1), C(n+1, 2), \ldots, C(n+k-1, k), \ldots)$ є функція $F(t) = (1-t)^{-n}$ для $|t| < 1$, тобто покажемо, що

$$(1-t)^{-n} = \sum_{k=0}^{\infty} C(n+k-1,k)t^k. \tag{5.31}$$

Скористаємося методом математичної індукції. Для $n = 1$ дана послідовність перетворюється на $(1, 1, \ldots, 1, \ldots)$, а $F(t) = (1-t)^{-1}$. Якщо $|t|<1$, то за формулою суми нескінченно спадної геометричної прогресії матимемо

$$1 + t + t^2 + \ldots + t^k + \ldots = (1-t)^{-1}. \tag{5.32}$$

Припустімо, що наше твердження виконується для $n = m$ $(m \geq 1)$, і розглянемо вираз

$$(1-t)^{-(m+1)} = (1-t)^{-m}(1-t)^{-1} = \big(C(m-1, 0) + C(m, 1)t + C(m+1, 2)t^2 + \ldots$$
$$\ldots + C(m+k-1, k)t^k + \ldots\big)\,(1 + t + t^2 + \ldots + t^k + \ldots).$$

Визначимо коефіцієнт при t^k в правій частині:

$$C(m-1, 0) + C(m, 1) + C(m+1, 2) + \ldots + C(m+k-1, k) =$$
$$= C(m+k, k).$$

Останню рівність неважко об'рунтувати, якщо замінити в ній $C(m-1, 0)$ на $C(m, 0)$ і скористатися тотожністю 2) з (5.8).

Отже,

$$(1-t)^{-(m+1)} = C(m, 0) + C(m+1, 1)t + C(m+2, 2)t^2 + \ldots$$
$$\ldots + C(m+k, k)t^k + \ldots,$$

що й доводить потрібне твердження.

Вище ми показали, що генератрисою послідовності $C(n, k)$ є функція $F(t) = (1 + t)^n$. Розглянемо таку послідовність: $a_k = 0$ для $k = 0, 1, 2, \ldots, n-1$ і $a_k = C(k, n)$ для $k \geq n$. Її генератрисою буде функція

$$F(t) = \sum_{k=n}^{\infty} C(k,n)t^k.$$

Виконавши рівносильні перетворення цього виразу з урахуванням співвідношення (5.31), маємо

$$F(t) = \sum_{k=n}^{\infty} C(k,n)t^k = t^n \sum_{m=0}^{\infty} C(n+m,n)t^m = t^n \sum_{m=0}^{\infty} C(n+m,m)t^m = \frac{t^n}{(1+t)^{n+1}}$$

Часто пошук аналітичної форми генератриси для заданої послідовності є непростою й до певної міри творчою задачею. Наприклад, генератриси деяких послідовностей можна знайти за допомогою рекурентних співвідношень.

Розглянемо послідовність $\underline{C(n, k)}$ кількостей сполук з n по k з повтореннями $k = 0, 1, 2, \ldots$. Вище було встановлено (див. (5.13), що

$$\underline{C(n, k)} = \underline{C(n, k-1)} + \underline{C(n-1, k)},\ k = 1, 2, \ldots . \tag{5.33}$$

Шукану генератрису послідовності $(\underline{C(n, 0)}, \underline{C(n, 1)}, \ldots, \underline{C(n, k)}, \ldots)$ позначимо

$$F_n(t) = \sum_{k=0}^{\infty} \underline{C(n,k)}t^k.$$

Помножимо обидві частини співвідношень (5.33) відповідно на t^k $(k = 1, 2, \ldots)$ і додамо почленно отримані рівності. Матимемо

$$\sum_{k=1}^{\infty}\underline{C(n,k)}t^k = t\sum_{k=1}^{\infty}\underline{C(n,k-1)}t^{k-1} + \sum_{k=1}^{\infty}\underline{C(n-1,k)}t^k,$$

тобто $F_n(t) - 1 = tF_n(t) + (F_{n-1}(t) - 1)$.

Отже, отримаємо таке рекурентне співвідношення:

$$F_n(t) = F_{n-1}(t) / (1 - t).$$

Звідси

$$F_n(t) = F_1(t) / (1 - t)^{n-1}.$$

Оскільки $\underline{C(1,k)} = 1$ для всіх k, то генератриса $F_1(t)$ цієї послідовності дорівнює

$$F_1(t) = \sum_{k=0}^{\infty}\underline{C(1,k)}t^k = 1 + t + t^2 + \ldots + t^k + \ldots = (1-t)^{-1}.$$

Таким чином, $F_n(t) = (1 - t)^{-n}$.

Порівнявши отриманий результат зі співвідношенням (5.31), утретє дістанемо формулу для числа сполук з n по k з повтореннями: $\underline{C(n,k)} = C(n + k - 1, k)$.

Аналогічно, за допомогою рекурентного співвідношення (5.15), можна довести, що генератрисою послідовності чисел Фібоначчі є $F(t) = (1 - t - t^2)^{-1}$. Розвинувши цю функцію в ряд, знову отримаємо формулу (5.22) для визначення довільного числа Фібоначчі.

Іноді генератрису можна знайти, виходячи з деяких комбінаторних міркувань. Вище ми встановили, що коефіцієнти a_k добутку біномів з формули (5.26) відповідають різним сполукам множини $\{x_1, x_2, \ldots, x_n\}$. Для цієї самої множини розглянемо такий добуток:

$$\prod_{i=1}^{n}(1 + x_i t + x_i^2 t^2 + \ldots + x_i^{m_i} t^{m_i}) = \sum_{k=0}^{s} a_k t^k. \qquad (5.34)$$

де $s = m_1 + m_2 + \ldots + m_n$.

Коефіцієнт a_k складається з добутків елементів усіх сполук з n по k з повтореннями множини $\{x_1, x_2, \ldots, x_n\}$ за умови, що кожен елемент x_i може з'являтися в цих сполуках не більше ніж m_i разів $(i = 1, 2, \ldots, n)$.

Отже, генератриса послідовності кількостей сполук з n по k з повтореннями із зазначеними обмеженнями дорівнює

$$F(t) = \prod_{i=1}^{n}(1 + t + t^2 + \ldots + t^{m_i}). \qquad (5.35)$$

Якщо на деякий елемент x_i не накладено жодних обмежень і він може повторюватись у сполуках будь-яку кількість разів, то в останній формулі слід покласти $m_i = \infty$, і відповідний множник матиме вигляд

$$1 + t + t^2 + \ldots + t^k + \ldots = (1 - t)^{-1}.$$

Зокрема, якщо всі елементи множини $\{x_1, x_2, \ldots, x_n\}$ можуть зустрічатися у сполуках без обмежень, то генератриса для послідовності кількостей таких сполук матиме вигляд:

$$F(t) = (1 + t + t^2 + \ldots + t^k + \ldots)^n = (1 - t)^{-n}.$$

Коефіцієнт при t^k у розвиненні цієї функції дорівнює $\underline{C(n,k)}$. Водночас він дорівнює $C(n + k - 1, k)$ (див. (5.31)). Отже, знову маємо формулу (5.11) для числа сполук з n по k з повтореннями.

Нарешті, якщо на сполуки з повтореннями накладено обмеження, що елемент x_i має входити до їх складу не менше ніж l_i разів і не більше ніж m_i разів, то відповідний множник у формулі (5.35) матиме вигляд $(t^{l_i} + t^{l_i+1} + \ldots + t^{m_i})$. А умові, що деякий елемент має повторюватись у сполуках лише парну кількість разів, у генератрисі (5.35) відповідатиме множник $(1 + t^2 + t^4 + \ldots + t^{2k} + \ldots)$.

Приклад 5.12. Позначимо $K(3, k)$ — число сполук із трьох елементів x_1, x_2, x_3 по k з повтореннями за умови, що x_1 зустрічається не більше одного разу, x_2 — не менше двох і не більше чотирьох разів, а x_3 — два чи п'ять разів. Усім таким сполукам відповідатимуть коефіцієнти многочлена

$$P(t) = (1 + x_1 t)(x_2^2t^2 + x_2^3t^3 + x_2^4t^4)(x_3^2t^2 + x_3^5t^5) = x_2x_2x_3x_3\,t^4 +$$
$$+ (x_2x_2x_2x_3x_3 + x_1x_2x_2x_3x_3)t^5 + (x_2x_2x_2x_2x_3x_3 + x_1x_2x_2x_2x_3x_3)t^6 +$$
$$+ (x_1x_2x_2x_2x_2x_3x_3 + x_2x_2x_3x_3x_3x_3x_3)t^7 + (x_2x_2x_2x_3x_3x_3x_3x_3 + x_1x_2x_2x_3x_3x_3x_3x_3)t^8 +$$
$$+ (x_2x_2x_2x_2x_3x_3x_3x_3x_3 + x_1x_2x_2x_2x_3x_3x_3x_3x_3)t^9 + x_1x_2x_2x_2x_2x_3x_3x_3x_3x_3\,t^{10},$$

тобто генератриса послідовності $(K(3, 0), K(3, 1), K(3, 2), \ldots, K(3, k), \ldots)$ дорівнює

$$F(t) = (1 + t)(t^2 + t^3 + t^4)(t^2 + t^5) = t^4 + 2t^5 + 2t^6 + 2t^7 + 2t^8 + 2t^9 + t^{10}.$$

Звідси матимемо

$K(3, 5) = K(3, 6) = K(3, 7) = K(3, 8) = K(3, 9) = 2$,

$K(3, 4) = K(3, 10) = 1$

$K(3, k) = 0$ для $k = 0, 1, 2, 3$ та $k > 10$.

Позначимо через $A_{n,k}$ кількість невід'ємних цілих розв'язків рівняння $x_1 + x_2 + \ldots + x_k = n$. Із результатів розділу 5.4 та формули (5.31) випливає, що функція $F(t) = (1 - t)^{-n}$ є генератрисою послідовності $A_{n,k}$, $n = 0, 1, 2, \ldots$. Неважко обґрунтувати, що генератрисою послідовності кількостей невід'ємних цілих розв'язків рівняння $m_1x_1 + m_2x_2 + \ldots \ldots + m_kx_k = n$, де $m_1, m_2, \ldots, m_k$ — фіксовані натуральні числа, буде функція

$$F(t) = \left((1 - t^{m_1})(1 - t^{m_2})\cdots(1 - t^{m_k})\right)^{-1}.$$

Послідовність $(a_0, a_1, a_2, \ldots, a_k, \ldots)$ позначатимемо $(a_k)_{k=0,\infty}$.

Нехай S — множина всіх послідовностей $(a_k)_{k=0,\infty}$ дійсних чисел. Означимо на S операції ***додавання*** та ***множення*** (*композиції*, або *згортки*) і позначимо їх + і °. Якщо $\alpha = (a_k)_{k=0,\infty}$ і $\beta = (b_k)_{k=0,\infty}$, то

$$\alpha + \beta = (a_k + b_k)_{k=0,\infty}, \tag{5.36}$$

$$\alpha \circ \beta = \left(a_0b_0,\ a_0b_1 + a_1b_0,\ a_0b_2 + a_1b_1 + a_2b_0,\ \ldots, \sum_{i=0}^{k} a_i b_{k-i}, \ldots\right). \tag{5.37}$$

Легко перевірити, що обидві операції асоціативні й комутативні, а операція множення дистрибутивна відносно додавання. Нейтральними елементами (нулем і одиницею) для додавання та множення є відповідно послідовності $\mathbf{0} = (0, 0, \ldots, 0, \ldots)$ і $\mathbf{1} = (1, 0, \ldots, 0, \ldots)$.

Для послідовності α оберненим елементом стосовно додавання буде послідовність $-\alpha = (-a_k)_{k=0,\infty}$. Обернений елемент щодо множення існує для тих і лише тих послідовностей α, у яких $a_0 \neq 0$. Отже, алгебра $A = \langle S, \{+, \circ\}\rangle$ — це комутативне кільце з одиницею.

Позначимо множину всіх генератрис послідовностей з S через G, а відображення, що ставить у відповідність послідовності $\alpha \in S$ її генератрису $F(t)$, — через γ. Неважко довести, що γ — ізоморфізм алгебр $A = \langle S, \{+, \circ\}\rangle$ і $A' = \langle G, \{+, \cdot\}\rangle$. Отже, алгебра генератрис A' — це також комутативне кільце з одиницею (одиницею A' є $F(t) \equiv 1$).

Побудованим ізоморфізмом γ зручно користуватися для визначення генератрис послідовностей $\alpha + \beta$ й $\alpha \circ \beta$, коли генератриси α і β відомі: $\gamma(\alpha + \beta) = \gamma(\alpha) + \gamma(\beta)$ й $\gamma(\alpha \circ \beta) = \gamma(\alpha) \cdot \gamma(\beta)$. Першу з цих властивостей уже було використано кілька разів у розв'язуванні різних задач, а другу властивість було неявно використано для обґрунтування співвідношень (5.28) і (5.31).

Приклад 5.13. Застосуємо другу властивість ізоморфізму γ для визначення генератриси $S(t)$ послідовності сум $S_k = \sum_{i=0}^{k} a_i$, коли відома генератриса $F(t)$ послідовності $\alpha = (a_k)_{k=0,\infty}$. Нехай $\sigma = (S_k)_{k=0,\infty}$ і $\beta = (1, 1, \ldots, 1, \ldots)$; тоді $\sigma = \alpha \circ \beta$. Оскільки генератрисою послідовності β є $(1 - t)^{-1}$, то генератрисою послідовності σ буде функція $S(t) = F(t)(1 - t)^{-1}$.

Додатково можна ввести операцію множення послідовності $\alpha = (a_k)_{k=0,\infty}$ на дійсне число r, вважаючи, що $r \cdot \alpha = (r \cdot a_k)_{k=0,\infty}$. Тоді $\gamma(r \cdot \alpha) = r \cdot \gamma(\alpha)$.

Генератриси можна застосовувати також для розв'язування рекурентних співвідношень (як лінійних, так і нелінійних).

Розглянемо спочатку операцію зсуву послідовності. Якщо $\alpha = (a_k)_{k=0,\infty}$, то ***зсувом*** послідовності α на m позицій називатимемо послідовність $\beta = (b_j)_{j=0,\infty}$, де $b_j = a_{m+j}$, $j = 0, 1, 2, \ldots$; позначатимемо $\beta = \Delta_m\alpha$. Якщо $F(t)$ — генератриса послідовності α, то генератрисою послідовності $\beta = \Delta_m\alpha$ буде функція

$$G(t) = \frac{F(t) - b_{m-1}(t)}{t^m}, \ 45\ b_{m-1}(t) = \sum_{i=0}^{m-1} a_i t^i. \tag{5.38}$$

Розглянемо лінійне однорідне рекурентне співвідношення з постійними коефіцієнтами

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \ldots + a_k f(n-k).$$

Помножимо обидві його частини на t^{n-k}, $n = k, k + 1, \ldots$, і додамо отримані рівності. З урахуванням операції зсуву дістанемо співвідношення

$$\frac{F(t) - b_{k-1}(t)}{t^k} = \sum_{i=1}^{k} \frac{a_i (F(t) - b_{k-i-1}(t))}{t^{k-i}},$$

де $F(t)$ — шукана генератриса послідовності $(f(0), f(1), f(2), \ldots, f(n), \ldots)$,

$$b_m(t) = \sum_{i=0}^{m} f(i) t^i, m = 0, 1, 2, \ldots, k-1, \text{ і } b_{-1}(t) = 0.$$

Звідси

$$F(t) = \frac{b_{k-1}(t) - \sum_{i=1}^{k-1} a_i b_{k-i-1}(t) t^i}{1 - \sum_{i=1}^{k} a_i t^i}. \tag{5.39}$$

Розвинувши цю функцію в ряд Маклорена, іншим способом дістанемо розв'язок рекурентного співвідношення (5.17).

Приклад 5.14. Нехай задано рекурентне співвідношення $f_{n+3} = 7f_{n+1} - 6f_n$ з початковими умовами $f_0 = -2, f_1 = 1, f_2 = 3$.

Тоді $b_2(t) = -2 + t + 3t^2$, $b_1(t) = -2 + t$ і $b_0(t) = -2$. За формулою (5.39) визначимо генерагрису $F(t)$ послідовності $(f_m)_{m=0,\infty}$:

$$F(t) = \frac{17t^2 + t - 2}{6t^3 - 7t^2 + 1}.$$

Оскільки $6t^3 - 7t^2 + 1 = (t-1)(2t-1)(3t+1)$, то за допомогою відомого методу невизначених коефіцієнтів генератрису нашої послідовності можна подати у вигляді

$$F(t) = \frac{-4}{1-t} + \frac{11}{5}\cdot\frac{1}{1-2t} - \frac{1}{5}\cdot\frac{1}{1+3t}.$$

Для розвинення в ряд кожного з доданків правої частини скористаємось формулою (5.32). Дістанемо

$$F(t) = \sum_{m=0}^{\infty} \frac{1}{5}(-20 + 11\cdot 2^m - (-3)^m)\,t^m.$$

Отже, розв'язком даного рекурентного співвідношення є функція

$$g(m) = \sum_{m=0}^{\infty} \frac{1}{5}(-20 + 11\cdot 2^m - (-3)^m).$$

До генератрис можна застосовувати також операції диференціювання й інтегрування.

Нехай $F(t)$ — генератриса послідовності $\alpha = (a_k)_{k=0,\infty}$, тоді генератрисою послідовності $k\alpha$ буде функція $tF'(t)$, а генератрисою послідовності $k(k-1)\alpha$ — функція $t^2F''(t)$. З останніх тверджень можна вивести, що генератриса послідовності $k^2\alpha$ — це функція $G(t) = t^2F''(t) + tF'(t)$.

Розглянемо інтеграл

$$H(t) = \int_0^t F(z)dz = \int_0^t (a_0 + a_1 z + a_2 z^2 + \ldots + a_k z^k + \ldots)dz =$$

$$= a_0 t + \frac{a_1}{2}t^2 + \frac{a_2}{3}t^3 + \ldots + \frac{a_k}{k+1}t^{k+1} + \ldots .$$

Отже, функція $H(t)$ є генератрисою послідовності $(b_k)_{k=0,\infty}$, де $b_0 = 0$ і $b_k = \frac{a_{k-1}}{k}$ для $k > 0$.

Використовуючи ці результати для генератриси $F(t) = (1-t)^{-1}$ послідовності (1, 1, …, 1, …), дістанемо, що генератрисою послідовності натуральних чисел (1, 2, 3, …) буде функція $(1-t)^{-2}$, а генератрисою послідовності $\left(0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\right)$□ DC=: FVO $\ln(1-t)^{-1}$.

Окрім того, якщо послідовність $(a_k)_{k=0,\infty}$ збіжна, то $\lim_{k\to\infty} a_k = \lim_{t\to 1}(1-t)F(t)$.

На завершення зауважимо, що в математиці існує багато різних способів означення та побудови генератрис. У найзагальнішому випадку запроваджують так звану базову послідовність функцій дійсної (або комплексної) змінної $(\varphi_0(t), \varphi_1(t), \varphi_2(t), \ldots, \varphi_k(t), \ldots)$ і генератрисою числової послідовності $\alpha = (a_k)_{k=0,\infty}$ називають функцію $F(t) = \sum_{k=0}^{\infty} a_k \varphi_k(t)$. Найчастіше (як це й було зроблено вище) як базову розглядають послідовність степеневих функцій $\varphi_k(t) = t^k$, $k = 0, 1, 2, \ldots$. У комбінаторному аналізі часто використовують як базові також експоненційні функції $\varphi_k(t) = \frac{t^k}{k!}$, $k = 0, 1, 2, \ldots$.

5.7. Задачі і вправи

1. Визначити генератрису послідовності $(a_k)_{k=0,\infty}$ ($c \in R$, $n \in N$):

1) $a_k = 1$;	5) $a_k = ck$;	9) $a_k = 1/(k+1)$;
2) $a_k = c^k$;	6) $a_k = k + c$;	10) $a_k = 1/k!$;
3) $a_k = kc^k$;	7) $a_k = k^2$;	11) $a_k = c^k/k!$;
4) $a_k = k + 1$;	8) $a_k = k^n$;	12) $a_k = C(n+k, k)$.

2. Нехай $F(t)$ — генератриса послідовності $(a_k)_{k=0,\infty}$. Визначити генератрису $G(t)$ послідовності $(b_k)_{k=0,\infty}$, якщо

1) $b_k = a_{k+1}$;
2) $b_k = a_{k+2}$;
3) $b_k = a_{k+n}$;
4) $b_k = a_k + 1$;
5) $b_k = a_k + c$;
6) $b_k = ca_k$;
7) $b_k = ca_k + d$;
8) $b_k = 0$, $0 \le k \le n$, і $b_k = a_k$ для $k > n$;
9) $b_k = 0$, $0 \le k \le n-1$, і $b_k = a_{k-n}$ для $k \ge n$;

10) $b_0 = a_0$ і $b_{k+1} = a_{k+1} - a_k$ для $k \geq 0$;
11) $b_k = a_{k+1} - a_k$;
12) $b_k = a_0 + a_1 + a_2 + \ldots + a_k$;
13) $b_k = ka_k$;
14) $b_k = k^2 a_k$;
15) $b_k = a_k/(k+1)$;
16) $b_k = a_k k/(k+1)$,
$c, d \in R$, $n \in N$.

3. Нехай $F(t)$ — генератриса послідовності $\alpha = (a_k)_{k=0..\infty}$, а послідовності $\beta_n = (b_k^{(n)})_{k=0..\infty}$, $n = 0, 1, 2, \ldots$, означають такими співвідношеннями: $b_k^{(0)} = a_k$, $b_k^{(n)} = \sum_{l=0}^{k} b_l^{(n-1)}$, $n = 1, 2, \ldots$. Визначити генератрису послідовності β_n, $n = 0, 1, 2, \ldots$.

4. Визначити формулу загального члена послідовності $(a_k)_{k=0..\infty}$, генератрисою якої є функція $F(t)$:

1) $F(t) = c$;
2) $F(t) = t^n$;
3) $F(t) = 1/(1-t)$;
4) $F(t) = 1/(1-t)^2$;
5) $F(t) = 1/(1-t)^n$;
6) $F(t) = 1/(1-ct)$;
7) $F(t) = ct/(1-ct)^2$;
8) $F(t) = ct/(1-t)^2$;
9) $F(t) = t(1+t)/(1-t)^3$;
10) $F(t) = ct(1+ct)/(1-ct)^3$;
11) $F(t) = t^n/(1-t)^{n+1}$;
12) $F(t) = t^n/(1-t)^n$;
13) $F(t) = (1-t)^{1/2}$;
14) $F(t) = (1-4t)^{1/2}$;
15) $F(t) = (1+t)^n$;
16) $F(t) = (c+dt)^n$;
17) $F(t) = t^n(1-t)^n$;
18) $F(t) = (1+t^2/2)^{-n}$;
19) $F(t) = \ln(1+t)$;
20) $F(t) = -\ln(1-t)$;
21) $F(t) = -\ln(1-ct)$;
22) $F(t) = (\ln((1+ct)/(1-ct)))/2$;
23) $F(t) = \operatorname{arctg} t$;
24) $F(t) = \arcsin t$, $c, d \in R$, $n \in N$.

5. Визначити послідовність $(a_k)_{k=0..\infty}$, генератриса якої дорівнює $F(t) = (1 + t + t^2 + \ldots + t^{n-1})^2$.

6. За допомогою заданої тотожності для генератрис і операції множення послідовностей довести відповідну тотожність для біноміних коефіцієнтів:

$$\sum_{l=0}^{\lfloor k/2 \rfloor} (-1)^l C(n+k-2l-1, n-1) C(n,l) = C(n,k);$$

(а) $(1+t)^n = (1-t^2)^n(1-t)^{-n}$;

$$\sum_{l=0}^{\lfloor k/2 \rfloor} C(n, k-2l) C(n+l-1, l) = C(n+k-1, k);$$

(б) $(1+t)^n(1-t^2)^{-n} = (1-t)^{-n}$;

$$\sum_{l=0}^{k} C(n,l) C(m, k-l) = C(n+m, k);$$

(в) $(1+t)^n(1+t)^m = (1+t)^{n+m}$;

$$\sum_{l=0}^{k} C(n+l, n) C(m+k-l, m) = C(n+m+k+1, k);$$

(г) $(1-t)^{-1-n}(1-t)^{-1-m} = (1-t)^{-2-n-m}$;

$$\sum_{l=0}^{k} (-1)^{k-l} C(n,l) C(m+k-l-1, k-l) = C(n-m, k);$$

(д) $(1+t)^n(1+t)^{-m} = (1+t)^{n-m}$;

$$\sum_{l=0}^{2k} (-1)^l C(n+l, n) C(m+2k-l, n) = C(n+k, k);$$

(е) $(1-t)^{-1-n}(1+t)^{-1-n} = (1-t^2)^{-1-n}$.

7. За допомогою визначення відповідної тотожності для генератрис і операції множення послідовностей довести таку тотожність для біномних коефіцієнтів:

(а) $\sum_{k=0}^{n} (-1)^k (C(n,k))^2$ дорівнює $(-1)^{n/2} C(n, n/2)$, якщо n парне, і дорівнює 0, якщо n — непарне;

(б) $\sum_{k} (-1)^{n-k} C(n,k) C(m+k, m+1) = C(m, n-1)$;

(в) $2\sum_{k} (-1)^k C(n, 2k) C(n, 2m-2k) = C(2n, 2m) + (-1)^m C(n,m)$;

(г) $2\sum_{k} (-1)^k C(n+2k, n) C(n+2m-2k+1, n) = C(2n+2m+2, 2n+1)$.

8. Члени послідовності $\varphi = (f_k)_{k=0..\infty}$ задовольняють лінійне однорідне рекурентне співвідношення зі сталими коефіцієнтами другого порядку $f_{n+2} + bf_{n+1} + cf_n = 0$.

(а) Визначити генератрису послідовності φ;

(б) знайти розв'язок заданого рекурентного співвідношення за допомогою методу генератрис.

9. Члени послідовності $\alpha = (a_k)_{k=\overline{0,\infty}}$ задовольняють рекурентне співвідношення $a_k = a_0a_{k-1} + a_1a_{k-2} + a_2a_{k-3} + \ldots + a_{k-1}a_0$, $k = 1, 2, \ldots$; $a_0 = 1$.

(а) Визначити генератрису послідовності α;

(б) розв'язати задане рекурентне співвідношення.

10. Обчислити суму:

(а) $C(n, 1) + 2C(n, 2) + 3C(n, 3) + \ldots + nC(n, n)$;

(б) $C(n, 0) + 2C(n, 1) + 3C(n, 2) + \ldots + (n + 1)C(n, n)$;

(в) $C(n, 2) + 2C(n, 3) + 3C(n, 4) + \ldots + (n - 1)C(n, n)$;

(г) $C(n, 0) + 3C(n, 1) + 5C(n, 2) + \ldots + (2n + 1)C(n, n)$;

(д) $C(n, 0) - 2C(n, 1) + 3C(n, 2) - \ldots + (-1)^n(n + 1)C(n, n)$;

(е) $3C(n, 1) + 7C(n, 2) + 11C(n, 3) + \ldots + (4n - 1)C(n, n)$;

(є) $C(n, 0)/1 + C(n, 1)/2 + \ldots + C(n, n)/(n + 1)$;

(ж) $C(n, 0)/2 - C(n, 1)/3 + C(n, 2)/4 - \ldots + (-1)^nC(n, n)/(n + 2)$.

11. Розв'язати рекурентні співвідношення задачі 5.6.5 за допомогою методу генератрис.

12. Розв'язати лінійне неоднорідне рекурентне співвідношення зі сталими коефіцієнтами $f_{n+2} + 2f_{n+1} - 8f_n = 2^n$, $f_0 = 0$, $f_1 = 2/3$.

13. Знайти рекурентне співвідношення для членів послідовності $(a_k)_{k=\overline{0,\infty}}$ і розв'язати це співвідношення:

(а) a_k — кількість способів розбиття опуклого $(k + 2)$-кутника на трикутники діагоналями, що не перетинаються всередині многокутника;

(б) a_k — число способів розставити дужки у виразі $b_1{\circ}b_2{\circ}\ldots{\circ}b_{n+1}$ так, щоб отримати правильні вирази;

(в) a_k — число способів сполучити попарно $2k$ точок, розміщених на колі, усіма можливими хордами, що не перетинаються всередині кола;

(г) a_k — число частин, на які розбивають площину k попарно непаралельних прямих, жодні три з яких не перетинаються в одній точці.

14. Визначити генератрису послідовності $(a_n)_{n=\overline{0,\infty}}$:

(а) a_n — кількість невпорядкованих зображень числа n у вигляді суми попарно різних цілих невід'ємних доданків;

(б) a_n — кількість невпорядкованих зображень числа n у вигляді суми натуральних доданків;

(в) a_n — кількість невпорядкованих зображень числа n у вигляді суми попарно різних непарних натуральних доданків;

(г) a_n — кількість невпорядкованих зображень числа n у вигляді суми непарних натуральних доданків;

(д) a_n — кількість розв'язків рівняння $k_1x_1 + k_2x_2 + \ldots + k_mx_m = n$ у цілих невід'ємних числах.

15. Визначити генератрису та формулу загального члена послідовності $(a_n)_{n=\overline{0,\infty}}$:

(а) a_n — кількість розв'язків рівняння $x_1 + x_2 + \ldots + x_m = n$ у цілих невід'ємних числах;

(б) a_n — кількість розв'язків рівняння $x_1 + x_2 + \ldots + x_m = n$ у натуральних числах.

РОЗДІЛ 6

БУЛЕВІ ФУНКЦІЇ

У найрізноманітніших галузях людської діяльності досить часто виникають об'єкти, явища, процеси чи ситуації, які характеризуються двома якісно різними станами або значеннями. Приклади таких альтернативних станів у техніці: "замкнено — розімкнено", "заряджено — незаряджено", "збуджено — спокійно", "працює — не працює", "є сигнал — немає сигналу" тощо. Аналогічні пари значень у логіці: "істинно — хибно", "так — ні", в алгебрі: "еквівалентно — нееквівалентно", "рівносильно — нерівносильно" тощо. При моделюванні таких об'єктів в математиці один з альтернативних станів (значень) позначають 1, а другий — 0. Зрозуміло, що в цьому випадку символи 0 і 1 не є числами у звичайному розумінні.

6.1. Булеві функції

Двоелементну множину $B = \{0, 1\}$ називають ***двійковим***, або ***булевим***, ***алфавітом***. Змінні, що набувають значень у B, називають ***двійковими***, або ***булевими***, ***змінними***.

Алгебра, носієм якої є множина B, називається ***алгеброю логіки***.

Булевою функцією (***логічною функцією***, або ***функцією алгебри логіки***) від n змінних називається будь-яка n-арна операція на B, тобто функція типу $B^n \to B$.

Множину всіх булевих функцій позначимо P_2, а множину всіх булевих функцій від n змінних — $P_2(n)$.

Кожну булеву функцію $f(x_1, x_2, \ldots, x_n)$ від n змінних можна цілком задати її таблицею значень (табл. 6.1), яку називають ***таблицею істинності*** функції f. У лівій частині цієї таблиці виписано всі двійкові кортежі

довжини n. Кількість таких кортежів — 2^n. Зазвичай кортежі розміщують у таблиці в лексикографічному порядку, який для *двійкових кортежів* $(\sigma_1, \sigma_2, \ldots, \sigma_n) \in B^n$ довжини n збігається зі звичайним порядком для відповідних двійкових чисел $\sigma_1\sigma_2\ldots\sigma_n$ або номерів цих кортежів.

Таблиця 6.1

x_1	x_2	...	x_{n-1}	x_n	$f(x_1, x_2, \ldots, x_{n-1}, x_n)$
0	0	...	0	0	$f(0, 0, \ldots, 0, 0)$
0	0	...	0	1	$f(0, 0, \ldots, 0, 1)$
0	0	...	1	0	$f(0, 0, \ldots, 1, 0)$
0	0	...	1	1	$f(0, 0, \ldots, 1, 1)$
...	...	...	...	...	....
1	1	...	1	0	$f(1, 1, \ldots, 1, 0)$
1	1	...	1	1	$f(1, 1, \ldots, 1, 1)$

Число $k = \sigma_1 \cdot 2^{n-1} + \sigma_2 \cdot 2^{n-2} + \ldots + \sigma_{n-1} \cdot 2 + \sigma_n$ називається ***номером*** кортежу $\alpha = (\sigma_1, \sigma_2, \ldots, \sigma_n)$; його позначають $\nu(\alpha)$. Наприклад, для $n = 4$ номером кортежу (1, 0, 0, 1) є число $1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2 + 1 = 9$.

Двійкові кортежі $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ часто записують у спрощеній формі $(\sigma_1\sigma_2\ldots\sigma_n)$ (або навіть $\sigma_1\sigma_2\ldots\sigma_n$). Далі будемо користуватися цією можливістю.

Вагою (або ***нормою***) двійкового вектора $\boldsymbol{a} = (a_1a_2\ldots a_n)$ називають число $a_1 + a_2 + \ldots + a_n$ (тобто кількість його одиничних координат); позначають $\|\boldsymbol{a}\|$.

Відстанню (*відстанню Хеммінга*) між двійковими векторами $\boldsymbol{a} = (a_1a_2\ldots a_n)$ і $\boldsymbol{b} = (b_1b_2\ldots b_n)$ називають число $\sum_{i=1}^{n}|a_i - b_i|$; позначають $\rho(\boldsymbol{a}, \boldsymbol{b})$. Двійкові вектори $\boldsymbol{a}, \boldsymbol{b} \in B^n$ називають ***сусідніми***, якщо $\rho(\boldsymbol{a}, \boldsymbol{b}) = 1$, і ***протилежними***, якщо $\rho(\boldsymbol{a}, \boldsymbol{b}) = n$.

Дві булеві функції від n змінних вважають ***різними***, якщо вони набувають різних значень принаймні на одному булевому кортежі. Отже, різні булеві функції мають різні стовпчики значень. Оскільки стовпчик значень для будь-якої булевої функції від n змінних є булевим кортежем довжини 2^n, і водночас кожен булевий кортеж довжини 2^n можна вважати стовпчиком значень якоїсь булевої функції від n змінних, то існує взаємно однозначна відповідність між $P_2(n)$

і множиною всіх двійкових кортежів довжини 2^n (тобто множиною B^{2^n}). Отже, має місце таке твердження.

Теорема 6.1. Кількість булевих функцій від n змінних дорівнює 2^{2^n}, тобто $|P_2(n)| = |B^{2^n}| = 2^{2^n}$.

Надалі, вважаючи, що двійкові вектори з B^n завжди розміщені в таблиці істинності в лексикографічному порядку, для задання функції f будемо використовувати лише стовпчик значень цієї функції, який записуватимемо у вигляді двійкового вектора $(b_0 b_1 b_2 \dots b_k)$, де $k = 2^n - 1$, називатимемо ***вектором значень*** функції f і позначатимемо $\boldsymbol{a}_f$. Координата b_i вектора $\boldsymbol{a}_f$ є значенням функції f на наборі значень змінних $(a_1, a_2, \dots, a_n)$ з номером i, $i = 0, 1, 2, \dots, 2^n - 1$.

Через N_f позначимо множину всіх двійкових векторів з B^n, на яких булева функція f набуває значення 1, тобто $N_f = \{\boldsymbol{a} \mid f(\boldsymbol{a}) = 1, \boldsymbol{a} \in B^n\}$.

Змінна x_i функції $f(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$ називається ***неістотною*** (***фіктивною***), якщо $f(x_1, x_2, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n) = f(x_1, x_2, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n)$ при будь-яких значеннях решти змінних. Інакше кажучи, зміна значення неістотної змінної x_i в будь-якому наборі значень аргументів $x_1, x_2, \dots, x_n$ не змінює значення функції f. У цьому разі функція $f(x_1, x_2, \dots, x_n)$ фактично не залежить від змінної x_i, й f можна вважати деякою функцією g від $n - 1$ змінної. Кажуть, що функцію g отримано з функції f шляхом ***вилучення фіктивної змінної***. При цьому функції f і g вважають ***рівними***. Співвідношення $f(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) = g(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ означає, що значення обох функцій збігаються для всіх значень змінних $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ незалежно від значення x_i.

Вилученню фіктивної змінної x_i з функції $f(x_1, x_2, \dots, x_n)$ відповідає викреслення в таблиці істинності функції f стовпчика змінної x_i, а також усіх рядків цієї таблиці виду $(\alpha_1, \alpha_2, \dots, \alpha_{i-1}, 1, \alpha_{i+1}, \dots, \alpha_n)$.

Процедура вилучення фіктивних змінних зрозуміла і природна, оскільки вони не впливають на значення функції й тому зайві. Однак інколи корисно вводити в деяку функцію фіктивні змінні, завдяки чому довільну функцію від n змінних можна зробити функцією від будь-якого більшого числа змінних. Отже, будь-яку скінченну сукупність функцій можна розглядати як множину функцій, що залежать від тих самих змінних.

Функції $f(x_1, x_2, \dots, x_n)$ і $g(x_1, x_2, \dots, x_m)$ називаються ***рівними*** (записують $f = g$), якщо функцію g можна одержати з функції f за допомогою вилучення і/або введення фіктивних змінних.

Функцію $f(x_1, x_2, \dots, x_n)$ називають ***невиродженою***, якщо вона не має фіктивних змінних.

6.1. Задачі і вправи

1. Знайти число двійкових векторів довжини n:

(а) k координат яких фіксовані;

(б) усі сусідні координати яких різні;

(в) серед сусідніх координат яких немає двох одиниць;

(г) координати яких утворюють неспадну послідовність;

(д) вага яких парна;

(е) вага яких непарна.

2. Знайти двійковий вектор найменшої довжини з номером k:

(а) $k = 21$; (б) $k = 1023$; (в) $k = 10001$; (г) $k = 2^7 - 2^4$.

3. Знайти двійковий вектор довжини n із номером k:

(а) $n = 5$, $k = 13$;

(б) $n = 10$, $k = 167$;

(в) $n = 10$, $k = 2^7 + 2^5$;

(г) $n = 10$, $k = 1023$;

(д) $n = m$, $k = 2^m - 1$;

(е) $n = m$, $k = 2^{m-1}$;

(є) $n = m$, $k = 2^{m-1} + 1$;

(ж) $n = m$, $k = 3 \cdot 2^{m-2} - 1$;

(з) $n = m$, $k = 7 \cdot 2^{m-3} - 1$.

4. Знайти кількість двійкових векторів $\boldsymbol{a}$ довжини n, номер $\nu(\boldsymbol{a})$ яких задовольняє такі умови:

(а) $\nu(\boldsymbol{a}) \geq 2^k$; (б) $2^{k-1} \leq \nu(\boldsymbol{a}) < 2^k$, $k = 1, 2, \dots, n$.

5. Довести, що двійкові вектори $\boldsymbol{a}$ і $\boldsymbol{b}$ довжини n протилежні тоді й тільки тоді, коли $\nu(\boldsymbol{a}) + \nu(\boldsymbol{b}) = 2^n - 1$.

6. Довести, що коли двійкові вектори $\boldsymbol{a}, \boldsymbol{b} \in B^n$ сусідні, то для деякого k ($k = 0, 1, 2, \dots, n - 1$) виконується умова $|\nu(\boldsymbol{a}) - \nu(\boldsymbol{b})| = 2^k$. Чи справедливе обернене твердження?

7. Для заданого двійкового вектора $\boldsymbol{a} \in B^n$ визначити кількість двійкових векторів довжини n:

(а) вага яких дорівнює $||\boldsymbol{a}||$;

(б) сусідніх з вектором $\boldsymbol{a}$;

(в) протилежних вектору $\boldsymbol{a}$;

(г) відстань яких від $\boldsymbol{a}$ більша k, $k = 1, 2, \dots, n$.

8. У множині B^n визначити кількість:

(а) невпорядкованих пар сусідніх векторів;

(б) упорядкованих пар протилежних векторів;

(в) упорядкованих пар векторів $\boldsymbol{a}$ і $\boldsymbol{b}$ таких, що $\rho(\boldsymbol{a}, \boldsymbol{b}) = k$, $k = 0, 1, 2, \dots, n$.

9. Означимо на множині B^n відношення R: для $\boldsymbol{a}, \boldsymbol{b} \in B^n$ вважаємо $\boldsymbol{a}\,R\,\boldsymbol{b}$ тоді й тільки тоді, коли $\|\boldsymbol{a}\| = \|\boldsymbol{b}\|$. Довести, що R — відношення еквівалентності. Визначити кількість класів еквівалентності за відношенням R і кількість елементів у кожному з цих класів.

10. Відображення f множини B^n в себе називається *монотонним*, якщо з $\nu(\boldsymbol{a}) \le \nu(\boldsymbol{b})$ випливає $\nu(f(\boldsymbol{a})) \le \nu(f(\boldsymbol{b}))$, $\boldsymbol{a}, \boldsymbol{b} \in B^n$. Визначити кількість усіх монотонних відображень множини B^n.

11. Знайти всі фіктивні змінні функції f, що задається вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (11110000)$;

(б) $\boldsymbol{a}_f = (1111000000001111)$.

Для функції f побудувати вектор значень невиродженої функції g, яка є рівною f.

12. Довести, що коли функція f має принаймні одну фіктивну змінну, то $|N_f|$ — парне число. Чи справедливим є обернене твердження?

6.2. Елементарні булеві функції

У теорії булевих функцій особливе місце займають функції однієї змінної — ***унарні*** ($n = 1$) і функції двох змінних — ***бінарні*** ($n = 2$), про які йтиметься нижче.

Існує $2^{2^1} = 4$ різних булевих функцій від однієї змінної (унарних булевих функцій, або унарних булевих операцій). Усі ці функції наведено в табл. 6.2.

Таблиця 6.2

x	ψ_0	ψ_1	ψ_2	ψ_3
0	0	0	1	1
1	0	1	0	1

Функції ψ_0 і ψ_3 — це константи 0 і 1; їх значення не залежать від значень змінної x, і тому вона неістотна. Функція ψ_1 "повторює" значення аргумента x $(\psi_1(x) = x)$ і називається ***тотожною***. Єдина нетривіальна функція — $\psi_2(x)$. Її називають функцією ***заперечення*** змінної x і позначають $\bar{x}$ тобто $\psi_2(x) = \bar{x}$ Функцію $\psi_2(x)$ називають також ***інверсією***, або ***функцією* НІ**, і позначають $\neg x$ або x'.

Кількість булевих функцій від двох змінних (бінарних булевих функцій, або бінарних булевих операцій) згідно з теоремою 6.1

дорівнює $2^{2^2} = 2^4 = 16$. Побудуємо зведену таблицю всіх цих функцій (табл. 6.3).

Таблиця 6.3

x	y	φ_0	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6	φ_7	φ_8	φ_9	φ_{10}	φ_{11}	φ_{12}	φ_{13}	φ_{14}	φ_{15}
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

Із 16 булевих функцій від двох змінних шість мають фіктивні змінні та зводяться до функцій від меншого числа змінних. Функції φ_0 і φ_{15} — це відповідно константи 0 і 1, тобто функції, обидві змінні яких неістотні. Функції φ_3 і φ_5 "повторюють" одну зі своїх змінних: $\varphi_3(x, y) = x$ і $\varphi_5(x, y) = y$, а функції φ_{10} і φ_{12} є запереченнями одної зі своїх змінних: $\varphi_{12}(x, y) = \bar{x}$ і $\varphi_{10}(x, y) = \bar{y}$.

Решта 10 функцій істотно залежать від обох своїх змінних. Для них використовують спеціальні назви і позначення за допомогою певних знаків бінарних операцій.

Функція $\varphi_1(x, y)$ називається ***кон'юнкцією*** (***логічним множенням***, ***функцією збігу*** чи ***логічним* І**) x і y; її позначають $x \wedge y$, $x \,\&\, y$, $\min(x, y)$ або xy (за аналогією зі звичайним числовим множенням знак кон'юнкції часто випускають і записують xy).

Функцію $\varphi_7(x, y)$ називають ***диз'юнкцією*** (***логічним додаванням***, ***з'єднувальним*** або ***логічним* АБО**) x і y; її позначають $x \vee y$, $\max(x, y)$ або $x + y$.

Функція $\varphi_6(x, y)$ — це ***додавання за модулем* 2**. Її позначають $x \oplus y$, $x \,\Delta\, y$ або $x \ne y$. Вона набуває значення 1, коли значення її аргументів різні, і значення 0, коли вони збігаються. Тому функцію φ_6 також називають функцією ***нерівнозначності***, ***альтернативою*** або ***розділовим* АБО**.

Функцію $\varphi_9(x, y)$ називають функцією ***еквівалентності***, або ***рівнозначності***, і позначають $x \sim y$, $x \leftrightarrow y$ або $x \equiv y$. Вона дорівнює 1, коли значення її аргументів однакові, і дорівнює 0, коли вони різні.

Функція $\varphi_{13}(x, y)$ носить назву ***імплікації*** і позначається $x \to y$ або $x \supset y$. Ці позначення читають так: "якщо x, то y" ("з x випливає y" або

"x імплікує y"). Звідси ще одна назва цієї функції — ***логічне слідування***.

Функцію $\varphi_{14}(x, y)$ називають ***штрихом Шеффера*** (***антикон'юнкцією***, або ***функцією* НІ-І**), позначають $x \mid y$; а функцію $\varphi_8(x, y)$ — ***стрілкою Пірса*** (***антидиз'юнкцією***, ***функцією* НІ-АБО**, ***функцією Даггера***, або ***функцією Вебба***), позначають $x \downarrow y$.

Інші функції: φ_2 — *антиімплікацію* ($\not\rightarrow$), φ_4 — *зворотну антиімплікацію* ($\not\leftarrow$) і φ_{11} — *зворотну імплікацію* ($\leftarrow$) використовують рідше, ніж вищезгадані.

Булеві константи 0 і 1 можна розглядати як виділені (фіксовані) елементи, тобто як нульарні операції алгебри логіки, або нульарні булеві функції.

Усі невироджені булеві функції від одної і двох змінних (унарні та бінарні булеві функції), а також нульарні булеві функції назвемо ***елементарними булевими функціями***.

Нижче буде доведено, що з елементарних булевих функцій за допомогою оператора суперпозиції S можна одержати всі булеві функції, тобто множина елементарних булевих функцій є системою твірних для алгебри $\langle P_2, \{S\}\rangle$.

6.2. Задачі і вправи

1. Побудувати таблицю істинності булевої функції f від трьох змінних x_1, x_2, x_3, що:

(а) набуває таке саме значення, як і більшість її змінних (*функція голосування*);

(б) набуває значення 1 на тих і лише тих наборах значень змінних, вага яких менша 2;

(в) на сусідніх наборах набуває різних значень;

(г) на протилежних наборах набуває однакових значень.

2. Чотирьом членам A_1, A_2, A_3, A_4 якогось комітету сформульовано такі умови відвідування засідань (принаймні одна з них має бути виконана): 1) у засіданні не беруть участі ні A_3, ні A_4, але має бути A_2; 2) у засіданні беруть участь A_1 і A_3, але відсутній A_2; 3) на засіданні повинні бути одночасно A_1 і A_4. Побудувати відповідну булеву функцію, яка описує правило відвідування засідань. З'ясувати, чи обов'язково повинен бути на засіданні комітету A_2, якщо в ньому не бере участі A_3.

3. По радіо можуть передаватися три рекламні повідомлення: $\boldsymbol{A}$, $\boldsymbol{B}$ і $\boldsymbol{C}$. Відомо, що на даний момент часу відбулася кожна з таких подій: 1) передано не більше ніж одне з повідомлень A і B; 2) повідомлення A могло бути передано в тому й тільки тому разі, якби були передані обидва повідомлення $\boldsymbol{B}$ і C; 3) передано принаймні одне з повідомлень A і C. Визначити, чи випливає звідси, що повідомлення $\boldsymbol{B}$ не передавалось, а повідомлення C — було передано.

4. За векторами значень булевих функцій f і g побудувати вектор значень функції h:

(а) $\boldsymbol{a}_f = (1011)$, $\boldsymbol{a}_g = (0101)$, $h(x_1, x_2) = f(f(x_1, x_2), g(x_1, x_2))$;

(б) $\boldsymbol{a}_f = (10101101)$, $\boldsymbol{a}_g = (1001)$, $h(x_1, x_2, x_3) = g(f(g(x_1, x_2), x_2, x_3), g(x_2, x_3))$.

5. Визначити, зі скількох елементів складається множина N_f для булевої функції f від n змінних, якщо f набуває:

(а) різних значень на протилежних наборах;

(б) різних значень на сусідніх наборах;

(в) значення 1 на тих і лише тих наборах, вага яких парна.

6. Визначити кількість булевих функцій f від n змінних, які задовольняють такі умови:

(а) f набуває фіксованого значення на заданому наборі $\boldsymbol{a} \in B^n$;

(б) f набуває фіксованого значення на заданих k наборах із множини B^n, $k = 0, 1, 2, \ldots, 2^n$;

(в) f набуває різних значень на протилежних наборах.

7. Визначити кількість булевих функцій від n змінних x_1, x_2, $\ldots x_n$, які не змінюються після перестановки (перейменування) у них будь-яких двох змінних x_i і x_j $(1 \le i < j \le n)$.

8. Довести, що $|N_f| = \|\boldsymbol{a}_f\|$.

6.3. Булеві функції і формули

Число булевих функцій від n змінних швидко зростає зі збільшенням n. Якщо для $n = 2$ воно дорівнює 16, то для $n = 3$ — 256, а для $n = 4$ — 65536. Зі зростанням числа аргументів збільшується й ускладнюється таблиця істинності булевої функції. Тому замість табличного способу задання булевих функцій часто використовують інші способи. Один із них — це спосіб задання булевих функцій за допомогою оператора суперпозиції та відповідних формул.

Нехай H — деяка множина булевих функцій, тобто $H \subseteq P_2$.

Оператор ***суперпозиції*** та поняття відповідної йому формули означимо індуктивно:

а) вважаємо суперпозиціями всі функції з H;

б) нехай задано сукупність S_1, S_2, $\ldots$, S_n, де S_i — або деяка булева змінна, або вже побудована раніше суперпозиція, $i = 1, 2, \ldots, n$, і f — n-арна булева функція з множини H. Тоді суперпозицією $S(S_1, S_2, \ldots, S_n, f)$ вважаємо вираз $f(S_1, S_2, \ldots, S_n)$.

Вираз, що описує цю суперпозицію та містить функціональні знаки і/або знаки операцій, круглі дужки й символи аргументів, називається ***формулою над*** H. Аргументи $S_1, S_2, \ldots, S_n$ суперпозиції називають ***підформулами***. Усі підформули нідформул $S_1, S_2, \ldots, S_n$ також називають підформулами основної формули. У суперпозиціях можна перейменовувати, переставляти і/або ототожнювати аргументи.

Розглянемо, наприклад, множину H, що складається з бінарних булевих функцій φ_1, φ_6, φ_{13} (відповідно кон'юнкції, додавання за модулем 2 й імплікації) й унарної булевої функції ψ_2 (заперечення). За допомогою цих функцій побудуємо кілька суперпозицій:

$$S_1(x, y) = \varphi_1(x, y);\ S_2(z) = \psi_2(z);$$

$$S_3(x, y, z) = S(S_1, S_2, \varphi_6) = \varphi_6(S_1(x, y), S_2(z)) = \varphi_6(\varphi_1(x, y), \psi_2(z));$$

$$S_4(x, y) = S(S_1, \psi_2) = \psi_2(S_1(x, y)) = \psi_2(\varphi_1(x, y));$$

$$S_5(x, y, z) = S(S_4, S_3, \varphi_{13}) = \varphi_{13}(S_4(x, y), S_3(x, y, z)) =$$
$$= \varphi_{13}(\psi_2(\varphi_1(x, y)), \varphi_6(\varphi_1(x, y), \psi_2(z))).$$

Якщо дозволити перейменовувати змінні в розглядуваних функціях і суперпозиціях, то, застосовуючи ті самі суперпозиції, можна замість S_5 дістати булеві функції від п'яти або чотирьох змінних, наприклад:

$$S_5'(x, y, z, u, v) = S(S_4(x, y), S_3(u, v, z), \varphi_{13}) =$$
$$= \varphi_{13}(\psi_2(\varphi_1(x, y)), \varphi_6(\varphi_1(u, v), \psi_2(z))),$$

$$S_5''(x, y, z, u) = S(S_4(x, y), S_3(u, y, z), \varphi_{13}) =$$
$$= \varphi_{13}(\psi_2(\varphi_1(x, y)), \varphi_6(\varphi_1(u, y), \psi_2(z))) \text{ тощо.}$$

Кінцеві функціональні вирази утворюють відповідні формули F_i для суперпозицій S_i, $i = 1, 2, \ldots, 5$. Назвемо їх ***функціональними***. Підформулами функціональної формули F_3 для S_3 будуть $\varphi_1(x, y)$ і $\psi_2(z)$, а підформулами формули F_5 для S_5 — $\psi_2(\varphi_1(x, y))$ і $\varphi_6(\varphi_1(x, y), \psi_2(z))$. Останні підформули, у свою чергу, складаються відповідно з підформул $\varphi_1(x, y)$ та $\varphi_1(x, y)$ і $\psi_2(z)$.

Усі отримані функціональні формули для суперпозицій S_i, $i = 1, 2, \ldots, 5$, можна переписати за допомогою введених вище спеціальних знаків бінарних операцій для булевих функцій φ_1, φ_6, φ_{13} і ψ_2:

$$F_1 = \varphi_1(x, y) = x \wedge y; \qquad F_2 = \psi_2(z) = \bar{z};$$

$$F_3 = \varphi_6(\varphi_1(x, y), \psi_2(z)) = (x \wedge y) \oplus \bar{z} \qquad F_4 = \psi_2(\varphi_1(x, y)) = (\overline{x \wedge y});$$

$$F_5 = \varphi_{13}(\psi_2(\varphi_1(x, y)), \varphi_6(\varphi_1(x, y), \psi_2(z))) = (\overline{x \wedge y}) \rightarrow ((x \wedge y) \oplus z).$$

На відміну від функціональних формул у лівих частинах наведених співвідношень формули у правих частинах умовно назвемо ***алгебричними***. Поряд із чисто функціональними або чисто алгебричними формулами для запису суперпозицій можна використовувати й так звані ***мішані формули***, наприклад для F_5 — формулу $\overline{\varphi_1(x, y)} \rightarrow$ $\rightarrow ((x \wedge y) \oplus \psi_2(z))$.

Під час побудови алгебричних формул уведення кожної операції супроводжується появою у формулі пари круглих дужок. Для спрощення запису алгебричних формул за аналогією зі звичайними алгебрами вводять поняття пріоритету, або старшинства операцій, в алгебрі логіки. Зокрема, вважають, що в першу чергу мають виконуватися операції заперечення, відтак кон'юнкції та, нарешті, диз'юнкції. Тоді дужки в алгебричних формулах можна використовувати тільки якщо треба змінити той порядок виконання операцій, який випливає з їх пріоритетів. Заперечення формули F записують за допомогою риски над усією формулою F; при цьому формулу F не обов'язково брати в дужки. Нарешті, нагадаємо, що знак кон'юнкції (логічного множення) між двома аргументами можна випускати.

Будь-якій суперпозиції, точніше відповідній їй формулі (функціональній, алгебричній або мішаній), відповідає певна булева функція $f \in P_2$. Цю відповідність також можна задати індуктивно:

а) формулі F для базової суперпозиції S_f, тобто функції $f(x_1, x_2, \ldots, x_n) \in H$, ставимо у відповідність саму цю функцію;

б) якщо формула F має вигляд $g(F_1, F_2, \ldots, F_m)$, де F_i — або формула (підформула) чи булева змінна x_{i_i}, і $g \in H$, то їй відповідатиме функція $g(f_1, f_2, \ldots, f_m)$, де f_i — відповідно або функція, що відповідає формулі F_i або тотожна функція $\iota(x_{i_i}) = x_{i_i}$, $i = 1, 2, \ldots, m$. Аргументами такої функції є аргументи або операнди даної формули.

Наведене індуктивне означення задає також правило (алгоритм) обчислення значення суперпозиції, а тим самим — і відповідних їй формули та булевої функції, для довільного набору значень їх аргументів. Коротко це правило можна сформулювати так: значення формули обчислюють тоді, коли обчислено значення всіх її підформул.

Запишемо, наприклад, послідовність кроків обчислення значення формули F_5 для набору значень $x = 1$, $y = 1$ і $z = 0$: $1 \wedge 1 = 1$, $\overline{0} = 1$, $1 \oplus 1 = 0$, $\overline{1} = 0$, $0 \to 1 = 1$, тобто $\overline{1 \wedge 1} \to ((1 \wedge 1) \oplus \overline{0}\,) = 1$, і формула F_5 та відповідна їй булева функція $f(x, y, z)$ на наборі значень (1, 1, 0) набувають значення 1. Аналогічно можна обчислити значення формули F_5 для всіх інших наборів значень її аргументів.

Словом, будь-яка формула F кожному набору α значень своїх аргументів (операндів), а отже, і кожному набору α значень змінних відповідної булевої функції f, ставить у відповідність значення цієї булевої функції f на наборі α. Отже, формула F може бути поряд із таблицею істинності способом задання й обчислення булевої функції f. Зокрема, за допомогою формули F, обчислюючи її значення на всіх можливих наборах значень аргументів, можна побудувати таблицю істинності булевої функції f.

Про формулу F, яка задає булеву функцію f, кажуть також, що вона ***реалізує***, або ***зображує***, функцію f. На відміну від табличного, спосіб задання булевих функцій за допомогою формул будемо називати ***аналітичним***.

Таблицю істинності для булевої функції $f(x, y, z)$, яку реалізує формула F_5, подано в табл. 6.4.

Таблиця 6.4

x	y	z	$f(x, y, z)$
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

Побудована відповідність між формулами і булевими функціями не є взаємно однозначною, бо вона не є ін'єктивною, і різні формули можуть реалізувати одну й ту саму булеву функцію. Наприклад, одну булеву функцію задають формули $F' = (x \wedge y) \vee (\overline{x} \wedge \overline{y}) \vee (\overline{y} \wedge \overline{z})$, $F'' = z \to (x \wedge y)$ і F_5 (у цьому можна переконатися за допомогою таблиць істинності).

6.3. Задачі і вправи

1. Визначити, чи є вираз A формулою над множиною функцій (операцій) H. У тому разі, коли A — не формула, чи можна, додавши дужки, перетворити її у формулу над H?

(а) $A = f_2(x, f_1(x, y))$, $H = \{f_1(x, y), f_2(x, y)\}$;

(б) $A = f_1(f_2(x, f_3(x, z), f_3(0, x)), f_3(f_2(0, y, z)))$, $H = \{0, f_1(x, y), f_2(x, y, z), \ f_3(x, y)\}$.

2. Виписати всі підформули формули A над множиною H:

(а) $A = f_2(f_1(x, f_3(x, y, z)), f_3(f_1(x, y), z, f_1(x, z)))$, $H = \{f_1(x, y), f_2(x, y), \ f_3(x, y, z)\}$;

(б) $A = ((x \to z) \wedge ((\neg x) \to (y \oplus z)))$, $H = \{\neg, \to, \wedge, \oplus\}$.

3. Довести, що будь-яку формулу A над множиною операцій $H = \{\neg, \wedge, \vee, \oplus, \sim, \to, |, \downarrow\}$ можна подати в одному й тільки в одному з таких виглядів: (a), $(\neg B)$, $(B \vee C)$, $(B \wedge C)$, $(B \oplus C)$, $(B \to C)$, $(B \sim C)$, $(B \mid C)$ або $(B \downarrow C)$, де a — змінна, а B і C — формули над H.

4. Використовуючи результат попередньої задачі, сформулювати правило, за яким для довільного виразу A за скінченне число кроків можна визначити, чи є A формулою над множиною операцій H.

5. Довести, що після заміни довільної підформули B формули A над множиною H на якусь формулу C над H отримаємо формулу над H.

6. Довести, що в результаті підстановки деякої формули B над множиною операцій H замість змінної a формули A над H знову отримаємо формулу над H.

7. Занумерувати послідовність виконання операцій у формулі й побудувати таблицю істинності булевої функції, заданої формулою:

(а) $(x \oplus (((x \to z) \vee (y \wedge x \wedge \overline{z}\,)) \mid (x \downarrow (x \vee \overline{z}))))$;

(б) $((((x \to y) \oplus \overline{z}) \downarrow z) \wedge (((x \vee ((z \mid x) \to x)) \vee z) \wedge x) \oplus ((\overline{y} \downarrow x) \sim z))$;

(в) $((((((x \to y) \to z) \oplus ((x \vee z) \to y)) \downarrow (((x \to \overline{y}) \oplus x) \downarrow x)) \mid z) \sim y)$.

8. За векторами значень булевих функцій f і g побудувати вектор значень функції h:

(а) $\boldsymbol{a}_f = (1101)$, $\boldsymbol{a}_g = (0111)$, $h(x_1, x_2) = f(f(x_1, x_2) \vee g(x_1, x_2), g(x_1, x_2) \to f(x_1, x_2))$;

(б) $\boldsymbol{a}_f = (11000101)$, $\boldsymbol{a}_g = (1011)$, $h(x_1, x_2, x_3, x_4) = f(g(x_1, x_3) \oplus (g(x_2, x_4) \to$ $\to (x_1 \vee x_3)), g(x_1 \oplus x_2, x_3 \to x_4), g(x_2, x_3))$.

9. Визначити всі бінарні булеві функції, які є результатом суперпозиції функцій множини D:

(а) $D = \{x \to y, x \wedge y\}$;

(б) $D = \{x \oplus y, x \oplus 1\}$;

(в) $D = \{f, g\}$, $a_f = (1, 1, 0, 1)$, $a_g = (10010110)$.

6.4. Основні тотожності алгебри логіки

Формули, які реалізують одну й ту саму булеву функцію, називаються ***рівносильними***. Рівносильність формул F_1 і F_2 позначають за допомогою знака рівності: $F_1 = F_2$ (іноді за допомогою знаків ~ або ≡). Отже, маємо

$$(\overline{x \wedge y}) \to ((x \wedge y) \oplus \bar{z}) = (x \wedge y) \vee (\bar{x} \wedge \bar{y}) \vee (\bar{y} \wedge \bar{z});$$

$$(\overline{x \wedge y}) \to ((x \wedge y) \oplus \bar{z}) = z \to (x \wedge y).$$

Неважко переконатися, що означена так рівносильність формул є відношенням еквівалентності на множині формул, тому часто рівносильні формули називають ***еквівалентними***.

Як для будь-яких двох формул визначити, рівносильні вони чи ні? Існує так званий ***стандартний метод***, який завжди дає змогу одержати вичерпну відповідь на це питання. Згідно з цим методом для кожної з даних формул слід побудувати таблицю істинності відповідної булевої функції, а відтак порівняти одержані таблиці.

Зокрема, за допомогою стандартного методу можна встановити такі співвідношення (у першому стовпчику записано рівносильності для алгебричних формул, а у другому — для відповідних функціональних формул):

$$\begin{array}{ll} x \oplus y = \bar{x}y \vee x\bar{y} & \varphi_6(x, y) = \varphi_7(\varphi_1(\psi_2(x), y), \varphi_1(x, \psi_2(y))), \\ x \sim y = \bar{x}\bar{y} \vee xy, & \varphi_9(x, y) = \varphi_7(\varphi_1(\psi_2(x), \psi_2(y)), \varphi_1(x, y)), \\ x \to y = \bar{x} \vee y, & \varphi_{13}(x, y) = \varphi_7(\psi_2(x), y), \\ x \mid y = \bar{x} \vee \bar{y} & \varphi_{14}(x, y) = \varphi_7(\psi_2(x), \psi_2(y)), \\ x \downarrow y = \bar{x}\bar{y} & \varphi_8(x, y) = \varphi_1(\psi_2(x), \psi_2(y)). \end{array} \qquad (6.1)$$

Крім того,

$$\begin{array}{l} \varphi_2(x, y) = \varphi_1(x, \psi_2(y)) = x\bar{y} \\ \varphi_4(x, y) = \varphi_1(\psi_2(x), y) = \bar{x}y, \\ \varphi_{11}(x, y) = \varphi_7(x, \psi_2(y)) = x \vee \bar{y} \end{array} \qquad (6.2)$$

Наведені рівносильності свідчать, що всі 10 невироджених бінарних булевих функцій можна за допомогою суперпозиції побудувати з булевих функцій кон'юнкції, диз'юнкції та заперечення. Інакше кажучи, для будь-якої формули над множиною елементарних булевих

функцій існує рівносильна їй формула, яка містить тільки знаки операцій $\wedge$, $\vee$ і $^-$. Отже, особливого значення набуває алгебра логіки $C = \langle B, \{\vee, \wedge, ^-\}\rangle$ типу (2, 2, 1). Іноді її записують як алгебру $\langle B, \{\vee, \wedge, ^-, 0, 1\}\rangle$ типу (2, 2, 1, 0, 0), відзначаючи в ній дві нульарні операції (фіксовані елементи) 0 і 1. Значну частину досліджень у теорії булевих функцій присвячено вивченню саме алгебри C.

Розглянемо основні властивості операцій алгебри C, подаючи їх у вигляді рівносильностей для відповідних формул. Формули алгебри C будемо називати ***булевими формулами*** (нагадаємо, що знак кон'юнкції $\wedge$ можна випускати).

1) $x(yz) = (xy)z$, $x \vee (y \vee z) = (x \vee y) \vee z$ (*асоціативність*),

2) $xy = yx$, $x \vee y = y \vee x$ (*комутативність*),

3) $(x \vee y)z = xz \vee yz$, $(xy) \vee z = (x \vee z)(y \vee z)$ (*дистрибутивність*),

4) $x \wedge 1 = x$, $x \vee 0 = x$, (*властивості*

5) $x \vee 1 = 1$, $x \wedge 0 = 0$, *елементів 0 та 1*), (6.3)

6) $x \vee \bar{x} = 1$, $x\bar{x} = 0$ (*властивості заперечення*),

7) $x \vee x = x$, $xx = x$ (*ідемпотентність*),

8) $\overline{xy} = \bar{x} \vee \bar{y}$, $\overline{x \vee y} = \bar{x}\,\bar{y}$ (*закони (правила) де Моргана*),

9) $\bar{\bar{x}} = x$ (*закон подвійного заперечення*).

Наведені рівносильності називають ***основними тотожностями*** (***еквівалентностями***) алгебри логіки C. Справедливість кожного з цих співвідношень можна встановити за допомогою стандартного методу. Безпосередньо з тотожностей 1–6 випливає, що алгебра $C = \langle B, \{\vee, \wedge, ^-, 0, 1\}\rangle$ *є булевою алгеброю*.

Зауважимо, що коли кожна з двох даних формул залежить від n змінних, то для перевірки їх рівносильності стандартним методом потрібно принаймні $2 \cdot 2^n$ кроків обчислень (для кожного з 2^n наборів значень аргументів), і, отже, для великих значень n практичне застосування цього методу може виявитися надто громіздким і незручним. Тому важливими задачами для довільної алгебри логіки є розробка ефективніших і зручніших методів установлення рівносильності даних формул і методів одержання формул, рівносильних даній.

Рівносильним перетворенням формули F називається таке перетворення, за допомогою якого з F отримуємо рівносильну їй формулу F'. Деякі з рівносильних перетворень для алгебри C випливають безпосередньо з її основних тотожностей.

Наприклад, з асоціативності та комутативності операцій кон'юнкції й диз'юнкції випливає можливість довільного порядку виконання дій при обчисленні кон'юнкції чи диз'юнкції будь-якого скінченного числа аргументів. Звідси випливає також, що можна записувати формули виду $x_1x_2\ldots x_n$ і $x_1 \vee x_2 \vee \ldots \vee x_n$ без дужок.

Далі, з властивості 5) елементів 0 і 1 випливає, що наявності хоча б одного нуля серед операндів формули виду $x_1x_2\ldots x_n$ достатньо для перетворення всього цього добутку в 0 і що коли хоча б один операнд формули виду $x_1 \vee x_2 \vee \ldots \vee x_n$ дорівнює 1, то і вся ця формула тотожно дорівнює 1. Ті самі властивості 4) і 5) дають змогу виключати з формул виду $x_1 \vee x_2 \vee \ldots \vee x_n$ усі операнди, що дорівнюють 0, а з формул виду $x_1x_2\ldots x_n$ — усі операнди, які дорівнюють 1. Очевидно, що рівносильними є й обернені перетворення: до будь-якої формули виду $x_1 \vee x_2 \vee \ldots \vee x_n$ можна додавати (диз'юнктивно) члени, які дорівнюють 0, а будь-яку формулу виду $x_1x_2\ldots x_n$ можна множити (кон'юнктивно) на члени, що дорівнюють 1.

Внаслідок тотожностей 7) виконуються рівності $x \vee x \vee \ldots \vee x = x$ та $xx\ldots x = x$, а з тотожності 9) випливає, що коли число заперечень над змінною x парне, то результат дорівнює просто x, а якщо непарне — $\bar{x}$

Отже, користуючись основними тотожностями алгебри логіки, можна одні булеві формули перетворювати в інші, рівносильні їм. Наприклад, за допомогою послідовного перетворення формул можна довести *правила поглинання*:

$$x \vee xy \overset{4}{=} (x \wedge 1) \vee (x \wedge y) \overset{3}{=} x(1 \vee y) \overset{5}{=} x \wedge 1 \overset{4}{=} x\,, \qquad (6.4)$$
$$x(x \vee y) \overset{4}{=} (x \vee 0)(x \vee y) \overset{3}{=} x \vee (0 \wedge y) \overset{5}{=} x \vee 0 \overset{4}{=} x.$$

Тут над кожним знаком рівності записано номер тієї основної тотожності, яку використано в цьому перетворенні.

6.4. Задачі і вправи

1. За допомогою стандартного методу з'ясувати, чи є рівносильними формули A і B:

(а) $A = ((x \vee y) \vee z) \to ((x \vee y) \wedge (x \vee z))$, $B = x \vee (y \sim z)$;

(б) $A = ((x \oplus y) \to (x \vee y)) \wedge ((\neg x \to y) \to (x \oplus y))$, $B = x \mid y$;

(в) $A = x \to (x \wedge y \to ((x \to y) \to y) \wedge z)$, $B = y \to (x \to z)$;

(г) $A = (x \to (y \oplus y)) \oplus y$, $B = x \sim y$;

(д) $A = ((x \to y) \wedge (y \to z)) \wedge (z \to x)$, $B = (x \vee y \vee z) \to (x \wedge y \wedge z)$;

(е) $A = (x \mid x) \mid ((x \mid y) \mid y)$, $B = x \vee y$.

2. За допомогою стандартного методу довести такі рівносильності:

(а) $x \oplus y = \bar{x}y \vee x\bar{y} = (x \vee y) \wedge (\bar{x} \vee \bar{y})$;

(б) $x \sim y = \bar{x}\bar{y} \vee xy = (x \vee \bar{y}) \wedge (\bar{x} \vee y)$;

(в) $x \to y = \bar{x} \vee y$;

(г) $x \mid y = \bar{x} \vee \bar{y} = \overline{x \wedge y}$;

(д) $x \downarrow y = \bar{x}\bar{y} = \overline{x \vee y}$.

3. Довести такі властивості операцій $\oplus$, $\to$, $\sim$, $\mid$, $\downarrow$:

(а) $x \oplus 0 = x$, $x \oplus 1 = \bar{x}$, $x \oplus x = 0$, $x \oplus \bar{x} = 1$;

(б) $x \to 0 = \bar{x}$, $x \to 1 = 1$, $x \to \bar{x} = \bar{x}$, $\bar{x} \to x = x$, $x \to x = 1$;

(в) $x \sim 0 = \bar{x}$, $x \sim 1 = x$, $x \sim x = 1$, $x \sim \bar{x} = 0$;

(г) $x \mid 0 = 1$, $x \mid 1 = \bar{x}$, $x \mid x = \bar{x}$, $x \mid \bar{x} = 1$;

(д) $x \downarrow 0 = \bar{x}$, $x \downarrow 1 = 0$, $x \downarrow x = \bar{x}$, $x \downarrow \bar{x} = 0$.

4. За допомогою стандартного методу довести основні тотожності алгебри логіки.

5. Довести, що:

(а) операції $\oplus$ і $\sim$ асоціативні;

(б) операції $\to$, $\mid$, $\downarrow$ неасоціативні;

(в) операції $\oplus$, $\sim$, $\mid$, $\downarrow$ комутативні;

(г) операція $\to$ некомутативна;

(д) операція $\vee$ дистрибутивна щодо $\sim$ і $\to$;

(е) операція $\vee$ недистрибутивна щодо $\oplus$, $\mid$, $\downarrow$.

6. Використовуючи основні тотожності алгебри логіки та результати попередніх задач, шляхом рівносильних перетворень довести рівносильність формул A і B:

(а) $A = (\bar{x} \to y) \to ((\bar{x} \wedge y) \sim (x \oplus y))$, $B = (\overline{x \wedge y} \to x) \to y$;

(б) $A = (x \wedge y \vee (\bar{x} \to y \wedge z)) \sim ((\bar{x} \to \bar{y}) \to z)$, $B = (x \to y) \oplus (y \oplus z)$.

7. Подати функцію f у вигляді формули над множиною операцій H:

(а) $f = \vee$, $H = \{\neg, \wedge\}$;

(б) $f = \to$, $H = \{\neg, \wedge\}$;

(в) $f = \vee$, $H = \{\neg, \to\}$;

(г) $f = \oplus$, $H = \{\neg, \vee\}$;

(д) $f = \oplus$, $H = \{\neg, \to\}$;

(е) $f = \vee$, $H = \{\mid\}$;

(є) $f = \oplus$, $H = \{\mid\}$;

(ж) $f = \to$, $H = \{\mid\}$;

(з) $f = \wedge$, $H = \{\downarrow\}$;

(и) $f = \to$, $H = \{\downarrow\}$;

(і) $f = \mid$, $H = \{\downarrow\}$;

(ї) $f = \sim$, $H = \{\wedge, \to\}$.

8. Довести, що функцію f не можна задати формулою над множиною операцій H:

(а) $f = \neg$, $H = \{\vee, \wedge, \to, \sim\}$;

(б) $f = \to$, $H = \{\vee, \wedge, \oplus\}$;

(в) $f = \sim$, $H = \{\vee, \wedge, \oplus\}$;

(г) $f = \oplus$, $H = \{\vee, \wedge, \to, \sim\}$.

9. Розв'язати рівняння:

(а) $(x \to \bar{x}) \to (\bar{x} \to x) = 0$;

(б) $(x \to (y \to z)) \to ((x \mid y) \mid z) = (x \mid z) \oplus y$.

10. Розв'язати систему рівнянь: $(x \vee y \vee z)(\bar{x} \vee \bar{y} \vee \bar{z}) = 0$, $x(\bar{y} \to \bar{z}) = 1$;

11. Скільки розв'язків має таке рівняння:

(а) $x_1 \vee x_2 \vee \ldots \vee x_n = 1$;

(б) $x_1 \oplus x_2 \oplus \ldots \oplus x_n = 0$;

(в) $x_1 \vee x_2 \vee \ldots \vee x_n = x_1 x_2 \ldots x_n$;

(г) $x_1 \oplus x_2 \oplus \ldots \oplus x_n = x_1 \vee x_2 \vee \ldots \vee x_n$;

(д) $x_1 \oplus x_2 \oplus \ldots \oplus x_n \oplus x_1 x_2 \ldots x_n = 1$;

(е) $\big((\ldots(x_1 \mid x_2) \mid x_3) \mid \ldots \mid x_{n-1}\big) \mid x_n = 0$.

12. Скільки розв'язків має система рівнянь:

(а) $x_1 \vee x_2 \vee \ldots \vee x_n = 1$, $x_1 \oplus x_2 \oplus \ldots \oplus x_n = 1$;

(б) $x_1 \vee x_2 \vee \ldots \vee x_n = x_1 x_2 \ldots x_n$, $x_1 \oplus x_2 \oplus \ldots \oplus x_n \oplus x_1 x_2 \ldots x_n = 1$;

(в) $x_1 \vee x_2 \vee \ldots \vee x_n = x_1 x_2 \ldots x_n$, $x_1 \oplus x_2 \oplus \ldots \oplus x_n = x_1 x_2 \ldots x_n$;

(г) $x_1 \bar{x}_2 \vee x_2 \bar{x}_3 \vee \ldots \vee x_{n-1} \bar{x}_n \vee x_n \bar{x}_1 = 1$, $x_1 \oplus x_2 \oplus \ldots \oplus x_n = x_1 x_2 \ldots x_n$.

13. Розв'язати нерівність:

(а) $(x \to \bar{x}) \to (\bar{x} \to x) < \bar{x} \vee (x \to \bar{x}) \vee \overline{x \sim x} \to \bar{x}$;

(б) $x \wedge y < x \sim y$.

14. Скільки розв'язків має нерівність:

(а) $x_1 \oplus x_2 \oplus \ldots \oplus x_n < x_1 x_2 \ldots x_n$;

(б) $1 \oplus x_1 \oplus x_1 x_2 \oplus x_1 x_2 x_3 \oplus \ldots \oplus x_1 x_2 \ldots x_n < x_1 \vee x_2 \vee \ldots \vee x_n$.

15. На множині Ф функцій, що залежать від змінних $x_1, x_2, \ldots, x_n$, означимо відношення $\leq$. Для функцій $f, g \in \Phi$ вважатимемо, що $f \leq g$, якщо для будь-якого набору $a_1, a_2, \ldots, a_n$ значень змінних для відповідних значень функцій виконується нерівність $f(a_1, a_2, \ldots, a_n) \leq g(a_1, a_2, \ldots, a_n)$ (вважаємо $0 < 1$). Довести, що $\leq$ — відношення часткового порядку на множині Ф.

16. Довести, що функції $f, g \in \Phi$ рівносильні, тобто $f = g$, тоді й тільки тоді, коли виконуються нерівності $f \leq g$ та $g \leq f$.

6.5. Розклад булевої функції за змінними

Розглядаючи відповідність між формулами та булевими функціями, зазначимо, що вона функціональна і всюди визначена, тобто є відображенням. Проте ця відповідність неін'єктивна, бо різні формули можуть зображувати одну й ту саму булеву функцію. Відтак з'ясуємо таке важливе питання: чи є сюр'єктивною розглядувана відповідність, тобто чи для кожної булевої функції з P_2 існує формула над певною множиною бу-

левих функцій, що зображує цю функцію? Відповідь на це питання дають наведені нижче теореми.

Введемо деякі позначення й означення: для булевої змінної x покладемо $x^0 = \bar{x}$ та $x^1 = x$. Тоді для $\alpha \in B$ виконується: $x^\alpha = 1$, якщо $x = \alpha$, і $x^\alpha = 0$, якщо $x \neq \alpha$.

Елементарною кон'юнкцією для набору змінних $x_1, x_2, \ldots, x_n$ назвемо кон'юнкцію деяких змінних цього набору або їх заперечень, у якій кожна змінна зустрічається не більше одного разу. Отже, елементарна кон'юнкція має вигляд $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \ldots x_{i_r}^{\alpha_r}$, $r \geq 1$; число r називають ***рангом*** елементарної кон'юнкції.

Диз'юнктивною нормальною формою (ДНФ) називається формула, яка має вигляд диз'юнкції елементарних кон'юнкцій. ***Диз'юнктивною нормальною формою булевої функції*** f називається ДНФ, яка зображує функцію f.

ДНФ називають ***досконалою (ДДНФ)***, якщо кожна її елементарна кон'юнкція містить усі змінні з набору.

Теорема 6.2 (про розклад булевої функції). Будь-яку булеву функцію $f(x_1, x_2, \ldots, x_n)$ можна подати у вигляді

$$f(x_1, x_2, \ldots, x_n) = \bigvee_{(\alpha_1, \alpha_2, \ldots, \alpha_m)} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_m^{\alpha_m} f(\alpha_1, \alpha_2, \ldots, \alpha_m, x_{m+1}, \ldots, x_n), \qquad (6.5)$$

де $1 \leq m \leq n$ і диз'юнкція здійснюється за всіма 2^m можливими наборами значень $\alpha_1, \alpha_2, \ldots, \alpha_m$ аргументів $x_1, x_2, \ldots, x_m$.

Доведення. Для доведення переконаємось, що булеві функції з лівої та правої частин рівності (6.5) набувають однакових значень при підстановці в них довільного набору $(\sigma_1, \sigma_2, \ldots, \sigma_n) \in B^n$ значень аргументів. Розглянемо праву частину. Оскільки $\sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \ldots \sigma_m^{\alpha_m} = 0$, якщо $(\alpha_1, \alpha_2, \ldots, \alpha_m) \neq (\sigma_1, \sigma_2, \ldots, \sigma_m)$, і $\sigma_1^{\sigma_1} \sigma_2^{\sigma_2} \ldots \sigma_m^{\sigma_m} = 1$, то серед усіх 2^m кон'юнкцій $\sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \ldots \sigma_m^{\alpha_m}$ правої частини лише одна дорівнюватиме 1, а саме та, в якій $\alpha_1 = \sigma_1$, $\alpha_2 = \sigma_2$, …, $\alpha_m = \sigma_m$. Отже,

$$f(\sigma_1, \sigma_2, \ldots, \sigma_n) = \bigvee_{(\alpha_1, \alpha_2, \ldots, \alpha_m)} \sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \ldots \sigma_m^{\alpha_m} f(\alpha_1, \alpha_2, \ldots, \alpha_m, x_{m+1}, \ldots, x_n) =$$

$$= \sigma_1^{\sigma_1} \sigma_2^{\sigma_2} \ldots \sigma_m^{\sigma_m} f(\sigma_1, \sigma_2, \ldots, \sigma_m, \sigma_{m+1}, \ldots, \sigma_n) = f(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

Теорему доведено.

Зображення (6.5) називається ***розкладом булевої функції*** f за змінними $x_1, x_2, \ldots, x_m$.

При $m = 1$ одержуємо розклад булевої функції за однією змінною x_1:

$$f(x_1, x_2, \ldots, x_n) = \overline{x}_1 f(0, x_2, x_3, \ldots, x_n) \vee x_1 f(1, x_2, x_3, \ldots, x_n). \qquad (6.6)$$

Зрозуміло, що аналогічно можна розкласти булеву функцію за будь-якою іншою змінною: $x_2, x_3, \ldots, x_n$.

Інший важливий випадок розкладу (6.5) — це розклад булевої функції за всіма n її змінними ($m = n$):

$$f(x_1, x_2, \ldots, x_n) = \bigvee_{(\alpha_1, \alpha_2, \ldots, \alpha_m)} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} f(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

Якщо булева функція $f(x_1, x_2, \ldots, x_n)$ не дорівнює тотожно 0, то в останній диз'юнкції залишаються тільки ті члени, для яких $f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1$, тобто можна записати

$$f(x_1, x_2, \ldots, x_n) = \bigvee_{\substack{(\alpha_1, \alpha_2, \ldots, \alpha_n) \\ f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1}} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}. \qquad (6.7)$$

Отже, отримано зображення булевої функції f у вигляді *досконалої диз'юнктивної нормальної форми* (ДДНФ).

Наприклад, булева функція $f(x, y, z)$, задана табл. 6.4, має таку ДДНФ:

$$f(x, y, z) = x^0y^0z^0 \vee x^0y^1z^0 \vee x^1y^0z^0 \vee x^1y^1z^0 \vee x^1y^1z^1 =$$
$$= \overline{x}\,\overline{y}\,\overline{z} \vee \overline{x}y\overline{z} \vee x\overline{y}\,\overline{z} \vee xy\overline{z} \vee xyz.$$

Зазначимо важливі *властивості ДДНФ*:

1) ДДНФ булевої функції f містить стільки елементарних кон'юнкцій, скільки одиниць у стовпчику значень її таблиці істинності;

2) кожному набору $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ значень змінних функції f, на якому вона дорівнює 1, у ДДНФ цієї функції відповідає єдина елементарна кон'юнкція всіх змінних $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$ така, що змінна x_i входить до неї із запереченням, якщо $\alpha_i = 0$, і без заперечення, якщо $\alpha_i = 1$.

Ці властивості можна вважати також ***правилами побудови ДДНФ*** для заданої булевої функції f ($f \neq 0$).

Отже, кожній булевій функції відповідає єдина (з точністю до порядку змінних в елементарних кон'юнкціях і порядку розміщення цих кон'юнкцій) ДДНФ.

З іншого боку, будь-якій ДДНФ вигляду $\bigvee_{i=1}^{k} x_1^{\alpha_{i_1}} x_2^{\alpha_{i_2}} \ldots x_n^{\alpha_{i_n}}$ однозначно відповідає певна n-арна булева функція φ. Стовпчик значень таблиці істинності функції φ містить k одиниць: вона набуває значення 1 тільки на наборах $(\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n})$ $i = 1, 2, \ldots, k$.

Отже, між множиною $P_2(n)$ всіх n-арних булевих функцій і множиною всіх формул, зображених у досконалій диз'юнктивній нормальній формі, існує взаємно однозначна відповідність. Єдина функція, що не має ДДНФ, — це функція, тотожно рівна 0.

Підсумком усього вищесказаного є така важлива теорема.

Теорема 6.3. Будь-яку булеву функцію можна зобразити булевою формулою, тобто формулою, операціями якої є кон'юнкція, диз'юнкція та заперечення.

Справді, для кожної булевої функції, окрім константи 0, таким зображенням є її ДДНФ. Константу 0 можна зобразити булевою формулою $x\overline{x}$ (див. (6.3), властивість 6).

Наслідок 6.3.1. Для будь-якої формули алгебри логіки існує рівносильна їй булева формула.

Справді, кожна формула F алгебри логіки реалізує певну булеву функцію f. Булева формула F', яка зображує функцію f, рівносильна формулі F.

6.5. Задачі і вправи

1. Знайти розклад булевої функції $f(x_1, x_2, \ldots, x_n)$, заданої своїм вектором значень $\boldsymbol{a}_f$, за всіма змінними із Z:

(а) $\boldsymbol{a}_f = (00100111)$, $Z = \{x_1\}$;

(б) $\boldsymbol{a}_f = (0101101010100011)$, $Z = \{x_1, x_3\}$.

2. Знайти розклад булевої функції $f(x_1, x_2, \ldots, x_n)$, заданої формулою A, за змінними із Z:

(а) $A = ((x_2 \to \overline{x}_1) \oplus (((x_3 \vee x_1) \sim \overline{x}_2) \to ((x_1 \mid x_3) \to x_1)) \oplus x_3)$, $Z = \{x_1\}$;

(б) $A = ((x_1 \mid ((x_2 \sim x_1) \to x_3)) \mid \overline{x}_4)$, $Z = \{x_2, x_4\}$.

3. Довести справедливість такого розкладу:

(а) $f(x_1, x_2, \ldots, x_n) = x_1(f(0, x_2, \ldots, x_n) \oplus f(1, x_2, \ldots, x_n)) \oplus f(0, x_2, \ldots, x_n)$;

(б) $f(x_1, x_2, \ldots, x_n) = (x_1 \vee (f(0, x_2, \ldots, x_n) \sim f(1, x_2, \ldots, x_n))) \sim f(1, x_2, \ldots, x_n)$.

4. Побудувати ДДНФ булевої функції $f(x_1, x_2, \ldots, x_n)$, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (0100)$;
(б) $\boldsymbol{a}_f = (10001001)$;
(в) $\boldsymbol{a}_f = (0001000101000110)$.

5. Побудувати ДДНФ булевої функції, заданої формулою A:

(а) $A = (x \oplus \overline{((y \to ((y \vee x) \mid y)) \sim x)} \wedge y)$;

(б) $A = (((z \to \overline{x \to (y \vee z)})) \mid ((z \sim x) \downarrow x)) \downarrow x)$.

6. Довести, що для будь-якої формули, яка не дорівнює тотожно 0, існує рівносильна їй диз'юнктивна нормальна форма.

7. За допомогою рівносильних перетворень побудувати ДНФ функції f, заданої формулою A:

(а) $A = (((x \oplus ((\overline{y \to z} \mid x) \sim y)) \to ((x \wedge y) \vee z)) \oplus z)$;

(б) $A = ((x \mid (y \to (((\overline{z \vee y} \sim x) \to x) \downarrow y))) \wedge z)$.

6.6. Алгебра формул і алгебра булевих функцій

Припустімо, що деяка множина булевих функцій H містить у собі елементарні булеві функції $\vee$, $\wedge$, $^-$, 0 і 1. Безпосередньо з означення поняття формули над H випливає, що коли F_1 і F_2 — формули над H, то вирази $F_1 \vee F_2$, $F_1 \wedge F_2$, $\overline{F_1}$, $F_1 \vee 0$, $F_1 \vee 1$, $F_1 \wedge 0$, $F_1 \wedge 1$ — також є формулами над H. Отже, система $D = \langle \Phi, \{\vee, \wedge, ^-, 0, 1\}\rangle$, де Φ — множина всіх формул алгебри логіки, є алгеброю, яку будемо називати ***алгеброю формул***.

Очевидно також, що коли F_1 і F_2 — булеві формули, то $F_1 \vee F_2$, $F_1 \wedge F_2$, $\overline{F_1}$, $F_1 \vee 0$, $F_1 \vee 1$, $F_1 \wedge 0$, $F_1 \wedge 1$ будуть булевими формулами. Це означає, що система $D_B = \langle \Phi_B, \{\vee, \wedge, ^-, 0, 1\}\rangle$, де Φ_B — множина всіх булевих формул, є підалгеброю алгебри формул D; D_B будемо називати ***алгеброю булевих формул***.

З означення рівносильності формул випливає справедливість таких правил перетворень в алгебрах формул. Нехай F_1 і F_2 — дві довільні рівносильні формули, тобто $F_1 = F_2$.

1. ***Правило підстановки.*** Нехай $x_1, x_2, \ldots, x_n$ — усі аргументи (операнди) формул F_1 і F_2, а $G_1, G_2, \ldots, G_n$ — довільний набір формул алгебри логіки. Позначимо через F_1' і F_2' формули, одержані з формул F_1 і F_2 після підстановки в них формул $G_1, G_2, \ldots, G_n$ замість усіх входжень аргументів $x_1, x_2, \ldots, x_n$ відповідно. Тоді F_1' і F_2' — рівносильні формули, тобто $F_1' = F_2'$.

2. ***Правило заміни.*** Припустімо, що F_1 — підформула формули F. Позначимо через F' формулу, одержану з формули F після заміни в ній підформули F_1 на рівносильну формулу F_2. Тоді $F = F'$.

Правило підстановки дає змогу узагальнити основні тотожності алгебри логіки (6.3) з булевих змінних x, y і z на довільні формули. Наприклад, для будь-яких формул A, B і C (зокрема, для булевих формул) виконуються такі рівносильності: $A(BC) = (AB)C$, $A \vee B = B \vee A$, $A \vee 1 = 1$, $A \vee \overline{A} = 1$, $\overline{AB} = \overline{A} \vee \overline{B}$ тощо.

Отже, алгебра формул $D = \langle \Phi, \{\vee, \wedge, ^-, 0, 1\}\rangle$ є булевою алгеброю. Булевою алгеброю є також алгебра $D_B = \langle \Phi_B, \{\vee, \wedge, ^-, 0, 1\}\rangle$.

Нехай булеву функцію f_1 задано формулою F_1, а булеву функцію f_2 — формулою F_2. Тоді згідно з правилом заміни для будь-якої пари формул F_1' і F_2' таких, що $F_1' = F_1$ і $F_2' = F_2$, формули $F_1' \vee F_2'$ і $F_1 \vee F_2$ рівносильні, тобто $F_1' \vee F_2' = F_1 \vee F_2$. Отже, для довільних формул F_1 і F_2, що зображують булеві функції f_1 і f_2, формула $F_1 \vee F_2$ однозначно визначає певну булеву функцію $f \in P_2$. Це дає змогу означити на множині булевих функцій P_2 похідну операцію диз'юнкції $\vee$, вважаючи $f_1 \vee f_2 = f$. Аналогічно можна означити на P_2 й операції кон'юнкції $\wedge$ та заперечення $^-$. Утворену в такий спосіб алгебру $L = \langle P_2, \{\vee, \wedge, ^-, 0, 1\}\rangle$ називатимемо ***алгеброю булевих функцій***.

Розглянемо відображення γ: $\Phi \to P_2$ і γ_B: $\Phi_B \to P_2$, що ставлять у відповідність кожній формулі з Φ або Φ_B булеву функцію, яку зображує ця формула. Безпосередньо з правил побудови алгебри L випливає, що γ та γ_B — гомоморфізми відповідно алгебри формул D і алгебри булевих формул D_B на алгебру булевих функцій L.

Аналогічно, користуючись наслідком із теореми 6.3 та правилом заміни, можна побудувати гомоморфне відображення η алгебри формул D в алгебру булевих формул D_B.

Позначимо через R уведене вище відношення рівносильності на множині формул (Φ та Φ_B). Із правила заміни випливає, що R — конгруенція для алгебр D і D_B. Отже, можна побудувати такі фактор-алгебри $D/R = \langle \Phi/R, \{\vee, \wedge, ^-, 0, 1\}\rangle$ і $D_B/R = \langle \Phi_B/R, \{\vee, \wedge, ^-, 0, 1\}\rangle$. Для кожної з фактор-множин Φ/R і Φ_B/R існують взаємно однозначні відображення δ і δ_B на множину всіх булевих функцій P_2: класу рівносильності фактор-множини формул ставимо у відповідність булеву функцію, яку зображують усі формули цього класу. Із тих самих правил побудови алгебри L

можна дійти висновку, що δ і δ_B — ізоморфні відображення відповідно алгебр D/R і D_B/R на алгебру L. Інакше кажучи, алгебри D/R, D_B/R і L ізоморфні між собою.

Оскільки алгебра формул D є булевою алгеброю, то гомоморфне відображення γ для кожної з аксіом булевої алгебри D породжує відповідне співвідношення для елементів алгебри L. Отже, алгебра булевих функцій L є булевою алгеброю, тому й ізоморфні їй фактор-алгебри D/R і D_B/R — також булеві алгебри.

Зазначені гомоморфні й ізоморфні відображення між алгебрами формул і алгеброю булевих функцій дуже часто призводять до неточностей та плутанини в термінології. Так, використовуючи терміни "рівносильні (еквівалентні) перетворення булевих функцій", "мінімізація чи оптимізація булевої функції", "спрощення булевої функції" тощо, мають на увазі перетворення, оптимізацію чи спрощення формул, які належать одному й тому самому класу рівносильних формул, що зображують одну булеву функцію, і перетворення, що не виводять за межі цього класу.

Підсумовуючи все вищесказане, можна побудувати таку діаграму відображень між усіма розглядуваними алгебрами і фактор-алгебрами (рис. 6.1).

$$\begin{tikzcd}
D \arrow[rr, "\psi"] \arrow[drr, "\gamma"] \arrow[dd, "\eta"'] & & D/R \arrow[d, leftrightarrow, "\delta"] \arrow[dd, bend left=60, leftrightarrow, "\theta"] \\
& & L \arrow[d, leftrightarrow, "\delta_B"] \\
D_B \arrow[urr, "\gamma_B"] \arrow[rr, "\psi_B"] & & D_B/R
\end{tikzcd}$$

Рис. 6.1

Тут η, γ, γ_B — гомоморфізми; ψ, ψ_B — канонічні гомоморфізми; δ, δ_B, θ — ізоморфізми.

6.6. Задачі і вправи

1. Довести, що відношення R рівносильності формул є відношенням еквівалентності на множині Φ формул над певною множиною операцій H.

2. Нехай рівносильні формули A і B містять змінну a. Довести, що в результаті підстановки деякої формули C замість усіх входжень змінної a у формули A і B отримаємо рівносильні формули A' і B'.

3. Нехай для формул B і C виконується рівність $B = C$, і B — підформула формули A. Довести, що після заміни у формулі A підформули B на C отримаємо формулу A_1, рівносильну A, тобто $A = A_1$.

4. Нехай R — відношення рівносильності на множині Φ формул над множиною операцій H. Довести, що існує взаємно однозначна відповідність між фактор-множиною Φ/R і певною підмножиною $P \subseteq P_2$. На основі цієї бієкції означити відповідні операції множини H для булевих функцій з P.

5. Дати обґрунтування діаграми на рис. 6.1, тобто довести, що η, γ, γ_B, ψ, ψ_B — гомоморфізми, а δ, δ_B, θ — ізоморфізми.

6.7. Канонічні форми булевих функцій

Неважко помітити, що правила підстановки і заміни є уточненням методу рівносильних перетворень, про який ішлося в розд. 4. Метод рівносильних перетворень — один з основних засобів доведення рівносильності формул у різних алгебрах, як правило, є потужнішим і ефективнішим, ніж описаний вище стандартний метод.

У різноманітних алгебрах, в яких є поняття рівносильності формул і рівносильних перетворень, виникає така класична проблема: для будь-яких заданих формул F_1 і F_2 з'ясувати, рівносильні вони чи ні, або чи існує ланцюг рівносильних перетворень, що переводять формулу F_1 у формулу F_2.

Найпоширеніший метод розв'язання цієї проблеми полягає в тому, що в класах рівносильних формул означають певних "стандартних представників", які дістали назву ***канонічної*** (або ***нормальної***) ***форми***. Відтак замість пошуків у "лабіринті" рівносильних перетворень "шляху" від F_1 до F_2 ставлять задачу зведення за допомогою рівносильних перетворень обох формул F_1 і F_2 до канонічної форми. Якщо одержані канонічні форми збігаються, то формули F_1 і F_2 рівносильні, інакше — нерівносильні.

При означенні канонічної форми в будь-якій алгебрі має бути виконано принаймні три такі умови або вимоги. По-перше, означена канонічна форма має існувати в кожному з класів еквівалентності. По-друге, вона має бути єдиною з точністю до неістотних відмінностей (наприклад, порядку розташування її елементів). Нарешті, має існувати конструктивна процедура (алгоритм), що дає змогу звести довільну формулу F з деякого класу еквівалентності рівносильних між собою формул до відповідної канонічної форми K, тобто шляхом рівносильних перетворень отримати з формули F формулу K.

Одну з канонічних форм — досконалу диз'юнктивну нормальну форму (ДДНФ) — було означено в розд. 5. Там було зазначено також, що кожній булевій функції, яка не дорівнює тотожно 0 (а тим самим — і кожному класу відповідних рівносильних формул), відповідає єдина ДДНФ. Отже, перші дві умови канонічної форми для ДДНФ виконуються. Алгоритм, що міститься в доведенні наступної теореми, забезпечує виконання третьої умови.

Теорема 6.4. Будь-яку булеву формулу F за допомогою певних рівносильних перетворень можна звести до ДДНФ.

Доведення. Доведемо теорему конструктивно. Сформулюємо набір правил, або алгоритм, застосовуючи який, довільну булеву формулу можна перетворити у відповідну їй ДДНФ. Головну роль у цьому алгоритмі відіграватимуть основні тотожності алгебри логіки (6.1) і правило заміни одних підформул на рівносильні їм інші підформули.

Формулюючи послідовність загальних правил, ілюструватимемо їх застосування на прикладі булевої формули

$$F = ((\overline{A \vee \overline{B}}) \vee (\overline{B \vee \overline{C}}) \vee (A \wedge C))) \wedge \overline{B}$$

1. Застосовувати правило 8) де Моргана доти, доки у формулі F заперечення стоятимуть лише над окремими змінними:

$$F = ((\overline{A} \wedge \overline{\overline{B}}) \vee (\overline{B \wedge \overline{C}} \vee (A \wedge C))) \wedge \overline{B} =$$
$$= ((\overline{A} \wedge B) \vee ((\overline{B} \vee \overline{\overline{C}}) \vee (A \wedge C))) \wedge \overline{B}.$$

2. Застосувати закон 9) подвійного заперечення так, щоб заперечення не зустрічалися двічі над однією й тією самою змінною:

$$F = ((\overline{A} \wedge B) \vee ((\overline{B} \vee C) \vee (A \wedge C))) \wedge \overline{B}.$$

3. Використати закон 3) дистрибутивності так, щоб формула складалася лише з диз'юнкції кон'юнкцій:

$$F = (\overline{A} \wedge B \wedge \overline{B}) \vee ((\overline{B} \vee C) \wedge \overline{B}) \vee (A \wedge C \wedge \overline{B}) =$$
$$= (\overline{A} \wedge B \wedge \overline{B}) \vee (\overline{B} \wedge \overline{B}) \vee (C \wedge \overline{B}) \vee (A \wedge C \wedge \overline{B}) .$$

4. До кожної кон'юнкції застосувати властивості 6) заперечення та властивість 7) ідемпотентності кон'юнкції так, щоб у кожну кон'юнкцію кожна змінна входила тільки один раз (із запереченням або без заперечення):

$$F = (\overline{A} \wedge 0) \vee \overline{B} \vee (C \wedge \overline{B}) \vee (A \wedge C \wedge \overline{B}).$$

5. Користуючись властивостями 4) елементів 0 і 1, вилучити з формули всі підформули, що містять 0 і 1.

$$F = \overline{B} \vee (C \wedge \overline{B}) \vee (A \wedge C \wedge \overline{B}).$$

6. Доданки (елементарні кон'юнкції) R, в які не входить певна змінна x, замінити двома доданками за правилом $R = (R \wedge x) \vee (R \wedge \overline{x})$, доки кожен доданок не міститиме всіх змінних даної формули. Справедливість цього правила, яке називають *правилом розщеплення*, випливає з таких рівностей:

$$(R \wedge x) \vee (R \wedge \overline{x}) = R \wedge (x \vee \overline{x}) = R \wedge 1 = R.$$

Тоді

$$F = (B \wedge A) \vee (\overline{B} \wedge \overline{A}) \vee (C \wedge \overline{B} \wedge A) \vee (C \wedge \overline{B} \wedge \overline{A}) \vee (A \wedge C \wedge \overline{B}) =$$
$$= (\overline{B} \wedge A \wedge C) \vee (\overline{B} \wedge A \wedge \overline{C}) \vee (\overline{B} \wedge \overline{A} \wedge C) \vee (B \wedge \overline{A} \wedge \overline{C}) \vee$$
$$\vee (C \wedge \overline{B} \wedge A) \vee (C \wedge \overline{B} \wedge \overline{A}) \vee (A \wedge C \wedge \overline{B}) .$$

7. Нарешті, застосувати властивість 7) ідемпотентності диз'юнкції так, щоб у формулу не входили дві однакові елементарні кон'юнкції. Дістанемо ДДНФ, рівносильну заданій формулі:

$$F = (\overline{B} \wedge A \wedge C) \vee (\overline{B} \wedge A \wedge \overline{C}) \vee (\overline{B} \wedge \overline{A} \wedge C) \vee (\overline{B} \wedge \overline{A} \wedge \overline{C}).$$

Додатково можна здійснити редагування одержаної ДДНФ: користуючись властивістю 2) комутативності кон'юнкції, упорядкувати за певним критерієм змінні в кожній елементарній кон'юнкції ДДНФ. Відтак, оскільки кожна кон'юнкція ДДНФ набуває значення 1 тільки на одному наборі значень змінних, то, користуючись комутативністю диз'юнкції, елементарні кон'юнкції ДДНФ можна впорядкувати згідно з відповідними булевими кортежами. Крім того, як і раніше, знак кон'юнкції можна не записувати. Остаточно дістанемо таку формулу:

$$F = \overline{ABC} \vee \overline{AB}C \vee A\overline{BC} \vee A\overline{B}C.$$

Якщо з формули F_1 за допомогою рівносильних перетворень можна одержати формулу F_2, то F_1 можна одержати з F_2, застосовуючи ті самі рівносильні перетворення у зворотному порядку. Інакше кажучи, будь-яке рівносильне перетворення оборотне. Це дає змогу довести таку теорему.

Теорема 6.5. Для будь-яких рівносильних булевих формул F_1 і F_2 існує послідовність рівносильних перетворень, яка переводить формулу F_1 у F_2 та навпаки.

Доведення. Можна вважати, що формули F_1 і F_2 містять ті самі змінні. В іншому разі в кожну з формул можна ввести відсутні (фіктивні) змінні за допомогою, наприклад, тотожності 6) $x \vee \bar{x} = 1$.

Із попередньої теореми випливає, що обидві булеві формули F_1 і F_2 можна звести до ДДНФ. Оскільки F_1 і F_2 рівносильні, то ці ДДНФ збігаються. Обертаючи послідовність перетворень формули F_2 в ДДНФ, одержимо шукану послідовність рівносильних перетворень $F_1 \Rightarrow$ ДДНФ $\Rightarrow F_2$. Аналогічно отримаємо обернене перетворення $F_2 \Rightarrow$ ДДНФ $\Rightarrow F_1$.

Для перевірки рівносильності формул, що містять інші операції, їх можна перетворити у булеві формули за допомогою співвідношень (6.1) та (6.2), відтак скористатись алгоритмом теореми 6.4.

Окрім ДДНФ в алгебрі булевих формул використовують ще одну канонічну форму — досконалу кон'юнктивну нормальну форму (ДКНФ).

Нехай $f(x_1, x_2, \ldots, x_n)$ — n-арна булева функція, що не дорівнює тотожно 1. Тоді, виходячи з ДДНФ функції $\bar{f}$ (6.7) і використовуючи закон 9) подвійного заперечення та закони 8) де Моргана, запишемо такий ланцюг рівносильних перетворень

$$f(x_1, x_2, \ldots, x_n) = \overline{\bar{f}}(x_1, x_2, \ldots, x_n) = \overline{\bigvee_{\substack{(\alpha_1, \alpha_2, \ldots, \alpha_n) \\ \bar{f}(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1}} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}} = \bigwedge_{\substack{(\alpha_1, \alpha_2, \ldots, \alpha_n) \\ f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0}} (x_1^{\bar{\alpha}_1} \vee x_2^{\bar{\alpha}_2} \vee \ldots \vee x_n^{\bar{\alpha}_n}). \quad (6.8)$$

Тут ми скористалися також тим, що функція $\bar{f}$ дорівнює 1 на тих і тільки тих двійкових кортежах, на яких f дорівнює 0, і, крім того, $\overline{x^{\bar{\alpha}}} = x^{\alpha}$.

Вираз у правій частині рівності (6.8) називається ***досконалою кон'юнктивною нормальною формою (ДКНФ)*** булевої функції f. Звідси можна дістати також правила побудови ДКНФ для n-арної булевої функції $f(x_1, x_2, \ldots, x_n)$.

1. ДКНФ булевої функції $f(x_1, x_2, \ldots, x_n)$ — це кон'юнкція всіх елементарних диз'юнкцій вигляду $x_1^{\bar{\alpha}_{i_1}} \vee x_2^{\bar{\alpha}_{i_2}} \vee \ldots \vee x_n^{\bar{\alpha}_{i_n}}$, $(\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n})$, $i = 1, 2, \ldots, k$, — усі двійкові кортежі, на яких функція f набуває значення 0.

2. В i-ту з цих диз'юнкцій змінна x_l входить із запереченням, якщо у відповідному двійковому кортежі координата $\alpha_{i_l} = 1$, і без заперечення, якщо $\alpha_{i_l} = 0$, $i = 1, 2, \ldots, k$; $l = 1, 2, \ldots, n$.

Наприклад, булева функція, задана табл. 6.4, має таку ДКНФ:

$$f(x, y, z) = (x \vee y \vee \bar{z}) \wedge (x \vee \bar{y} \vee \bar{z}) \wedge (\bar{x} \vee y \vee \bar{z}).$$

Безпосередньо з означення ДКНФ і правил її побудови можна довести, що для будь-якої булевої функції f, яка не дорівнює тотожно 1, існує ДКНФ, і вона єдина з точністю до порядку змінних у диз'юнкціях і порядку розташування самих диз'юнкцій. Справедливість цього твердження випливає також зі зв'язку між ДКНФ функції f і ДДНФ функції $\bar{f}$ (див. (6.8)). Цей зв'язок можна узагальнити й сформулювати так званий *принцип двоїстості* для булевих функцій.

6.7. Задачі і вправи

1. За допомогою рівносильних перетворень звести задані ДНФ до ДДНФ:

(а) $\bar{x}y \vee y \vee x$; (б) $\bar{x}y \vee \bar{y}z \vee x\bar{y}z \vee z$; (в) $\bar{z} \vee \bar{y}z \vee x\bar{y} \vee xz$.

2. За допомогою зведення до ДДНФ довести рівносильність формул A і B:

(а) $A = \overline{(\bar{x} \vee y)} \vee \overline{(\bar{y} \vee x)}$, $B = (x \vee y) \wedge (\bar{x} \vee \bar{y})$;

(б) $A = (\bar{x} \wedge \bar{z}) \vee (x \wedge y) \vee (x \wedge \bar{z})$, $B = x \wedge (\bar{z} \vee y \wedge z) \vee (\bar{x} \wedge \bar{z})$.

3. Нехай задано ДДНФ D_1 і D_2 булевих функцій $f(x_1, x_2, \ldots, x_n)$ і $g(x_1, x_2, \ldots, x_n)$. Описати процедуру побудови ДДНФ функції $h(x_1, x_2, \ldots, x_n)$, заданої формулою D:

(а) $D = \neg D_1$; (в) $D = D_1 \wedge D_2$; (д) $D = D_1 \to D_2$;

(б) $D = D_1 \vee D_2$; (г) $D = D_1 \oplus D_2$; (е) $D = D_1 \sim D_2$.

4. Описати процедуру побудови ДДНФ булевої функції, заданої формулою $C = A \oplus B$, якщо відомі ДДНФ D_1 і D_2 функцій, заданих формулами:

(а) $D_1 = A \wedge B$, $D_2 = A \vee B$; (б) $D_1 = A \to B$, $D_2 = B \to A$.

5. Довести, що коли в ДДНФ функції f замінити скрізь операцію $\vee$ на $\oplus$, то отримаємо формулу, що задає функцію f, тобто рівносильну формулу. Чи справедливе аналогічне твердження для довільної ДНФ функції f?

6. Побудувати ДКНФ булевої функції $f(x_1, x_2, \ldots, x_n)$, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (1100)$;

(б) $\boldsymbol{a}_f = (01010011)$;

(в) $\boldsymbol{a}_f = (1101011101001110)$.

7. Побудувати ДКНФ булевої функції від трьох змінних, яка набуває такого самого значення, як більшість її змінних.

8. Нехай формулу A записано в ДКНФ. Побудуємо формулу B так:

1) випишемо кон'юнкцію всіх елементарних диз'юнкцій, які не входять до A;

2) замінимо $\wedge$ на $\vee$, $\vee$ — на $\wedge$, x_i — на $\overline{x}_i$, $\overline{x}_i$ — на x_i.

Довести, що формула B — це ДДНФ формули A.

9. Нехай задано ДДНФ формули A. Описати процедуру побудови ДКНФ формули A.

10. Описати процедуру побудови за ДДНФ формули A і ДДНФ формули B:

(а) ДКНФ і ДДНФ формули $(\overline{A})$;

(б) ДКНФ і ДДНФ формули $(A \vee B)$;

(в) ДКНФ і ДДНФ формули $(A \wedge B)$;

(г) ДКНФ і ДДНФ формули $(A \to B)$.

6.8. Принцип двоїстості

Булева функція $g(x_1, x_2, \ldots, x_n)$ називається ***двоїстою*** до булевої функції $f(x_1, x_2, \ldots, x_n)$, якщо $g(x_1, x_2, \ldots, x_n) = \overline{f}\,(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)$. Функцію, двоїсту до булевої функції f, позначають f^*. Отже, $f^*(x_1, x_2, \ldots, x_n) = \overline{f}(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)$.

Таблицю істинності функції f^* можна одержати з таблиці істинності функції f, якщо в останній інвертувати (тобто замінити 0 на 1, 1 на 0) усі значення аргументів функції, а також усі значення в стовпчику значень функції. Із цієї властивості випливає, зокрема, що двоїстою до функції f^* є сама функція f, тобто $(f^*)^* = f$.

Булева функція, двоїста до самої себе, тобто функція f, для якої виконується $f^* = f$, називається ***самодвоїстою***. Самодвоїста n-арна функція на протилежних наборах $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ і $(\overline{\alpha}_1, \overline{\alpha}_2, \ldots, \overline{\alpha}_n)$ набуває протилежних значень.

Неважко переконатися, що диз'юнкція двоїста до кон'юнкції, нульарна функція-константа 1 двоїста до функції-константи 0, а функція заперечення самодвоїста. Ще один приклад самодвоїстої функції — $m(x, y, z) = xy \vee xz \vee yz$ (*функція медіана*, або *функція голосування*).

Користуючись означенням двоїстості, можна довести твердження, яке називають принципом двоїстості.

Теорема 6.6 (***принцип двоїстості***). Якщо формула $F = S(f_1, f_2, \ldots, f_k, \psi) = \psi(f_1, f_2, \ldots, f_k)$ реалізує булеву функцію $\varphi(x_1, x_2, \ldots, x_n)$, то формула $F^* = S(f_1^*, f_2^*, \ldots, f_k^*, \psi^*) = \psi^*(f_1^*, f_2^*, \ldots, f_k^*)$, одержана з F заміною

функцій $f_1, f_2, \ldots, f_k$, ψ на двоїсті функції $f_1^*, f_2^*, \ldots, f_k^*$, ψ^*, реалізує булеву функцію $\varphi^*(x_1, x_2, \ldots, x_n)$, двоїсту до функції $\varphi(x_1, x_2, \ldots, x_n)$. Тут $x_1, x_2, \ldots, x_n$ — операнди формули F (усі аргументи булевих функцій $f_1, f_2, \ldots, f_k$).

Доведення. Справді, нехай формула $F = S(f_1, f_2, \ldots, f_k, \psi) = \psi(f_1, f_2, \ldots, f_k) = \psi(f_1(x_{11}, x_{12}, \ldots, x_{1l_1}), f_2(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, f_k(x_{k1}, x_{k2}, \ldots, x_{kl_k}))$ реалізує булеву функцію $\varphi(x_1, x_2, \ldots, x_n)$. Тоді $F^* = \psi^*(f_1^*(x_{11}, x_{12}, \ldots, x_{1l_1}), f_2^*(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, f_k^*(x_{k1}, x_{k2}, \ldots, x_{kl_k})) = \overline{\psi}\,(\overline{f_1^*}\,(x_{11}, x_{12}, \ldots, x_{1l_1}), \overline{f_2^*}\,(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, \overline{f_k^*}\,(x_{k1}, x_{k2}, \ldots, x_{kl_k})) = \overline{\psi}\,(\overline{f}_1(\overline{x}_{11}, \overline{x}_{12}, \ldots, \overline{x}_{1l_1}), \overline{f}_2(\overline{x}_{21}, \overline{x}_{22}, \ldots, \overline{x}_{2l_2}), \ldots, \overline{f}_k(\overline{x}_{k1}, \overline{x}_{k2}, \ldots, \overline{x}_{kl_k})) = \overline{\psi}\,(f_1(\overline{x}_{11}, \overline{x}_{12}, \ldots, \overline{x}_{1l_1}), f_2(\overline{x}_{21}, \overline{x}_{22}, \ldots, \overline{x}_{2l_2}), \ldots, f_k(\overline{x}_{k1}, \overline{x}_{k2}, \ldots, \overline{x}_{kl_k}))$.

Отже, формула F^*, яку природно називати формулою *двоїстою* до F, реалізує булеву функцію $\overline{\varphi}(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)$, але $\overline{\varphi}(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n) = \varphi^*(x_1, x_2, \ldots, x_n)$.

У термінах булевих формул принцип двоїстості можна сформулювати так: якщо в булевій формулі F, яка зображує булеву функцію f, замінити всі операції кон'юнкції на операції диз'юнкції, операції диз'юнкції — на операції кон'юнкції, усі нульарні операції 0 замінити на 1, а всі 1 — на 0, то одержимо формулу F^*, яка зображує булеву функцію f^*, двоїсту до функції f.

Якщо булеві функції f і g рівні між собою, то рівні й відповідні їм двоїсті функції f^* і g^*, і навпаки. Це означає, що формули F_1^* і F_2^*, двоїсті до рівносильних формул F_1 і F_2, також рівносильні.

Останній висновок дає змогу для будь-якої тотожності $F_1 = F_2$ в алгебрі формул автоматично одержувати нову тотожність $F_1^* = F_2^*$ для відповідних двоїстих формул. Такі тотожності можна умовно назвати *двоїстими*. Наприклад, пари основних тотожностей 1)–8) (6.3) алгебри логіки двоїсті. Двоїстими є також правила поглинання (6.4). Для алгебри булевих формул принцип двоїстості іноді називають ще *узагальненим законом де Моргана*.

Якщо M — деяка множина функцій з P_2, то через M^* позначимо множину всіх функцій, двоїстих до функцій з M. Множина M^* називається

двоїстою до множини M. Якщо $M^* = M$, то множина M називається ***самодвоїстою***.

6.8. Задачі і вправи

1. Побудувати таблицю істинності та вектор значень функції g, двоїстої до булевої функції f, яку задано вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (0011)$;

(б) $\boldsymbol{a}_f = (00000001)$;

(в) $\boldsymbol{a}_f = (1010100011000001)$.

2. Перевірити, чи є функція f, задана формулою A, двоїстою до функції g, заданої формулою B:

(а) $A = (x \to y)$, $B = (\overline{x} \wedge y)$;

(б) $A = x \oplus y \oplus z$, $B = x \oplus y \oplus z$;

(в) $A = x \wedge y \vee z$, $B = x \wedge (y \vee z)$;

(г) $A = xy \oplus xz \oplus yz$, $B = xy \vee xz \vee yz$.

3. Нехай $\boldsymbol{a}_f = (a_0, a_1, a_2, \ldots, a_k)$ — вектор значень булевої функції f від n змінних. Довести, що вектор значень двоїстої функції f^* має вигляд $(\overline{a_k}, \overline{a_{k-1}}, \ldots, \overline{a_1}, \overline{a_0})$, де $k = 2^n - 1$.

4. Нехай g — функція, двоїста до булевої функції f від n змінних. Довести що:

(а) $|N_f| + |N_g| = 2^n$;

(б) $|N_f| - |N_g|$ — число парне;

(в) $\rho(\boldsymbol{a}_f, \boldsymbol{a}_g)$ — число парне.

5. Довести, що дані функції двоїсті одна до одної:

(а) заперечення і заперечення;

(б) диз'юнкція та кон'юнкція;

(в) функція $\oplus$ і функція $\sim$;

(г) функція $|$ і функція $\downarrow$.

6. Довести, що формули A і B рівносильні тоді й тільки тоді, коли рівносильні формули A^* і B^*.

6.9. Проблема повноти для алгебри булевих функцій

Аналітичний спосіб задання булевих функцій є, безумовно, зручніший і компактніший, ніж табличний. Джерелом ефективності аналітичного способу є можливість послідовної побудови й обчислення булевої функції за допомогою суперпозиції інших (простіших) булевих функцій із деякої заданої сукупності функцій $H \subseteq P_2$. Однак у разі його застосування завжди виникає питання: чи будь-яку булеву функцію з P_2 можна подати формулою над заданою базовою сукупністю булевих функцій H? Інакше кажучи, чи є задана сукупність булевих функцій H системою твірних алгебри $\langle P_2, \{S\}\rangle$, або чи виконується рівність $[H]_S = P_2$? Отже, постає ***проблема повноти*** для певної множини H стосовно алгебри $\langle P_2, \{S\}\rangle$.

Будь-яку систему твірних Σ алгебри $\langle P_2, \{S\}\rangle$ традиційно називають ***функціонально повною системою*** булевих функцій. Отже, система булевих функцій Σ називається ***функціонально повною***, якщо $[\Sigma]_S = P_2$, тобто коли довільну булеву функцію можна зобразити формулою над Σ.

Із теореми 6.3 випливає, наприклад, що система $\Sigma_0 = \{\vee, \wedge, ^{-}\}$ функціонально повна. Зрозуміло, що функціонально повною буде й будь-яка сукупність Σ булевих функцій, з яких за допомогою операції суперпозиції можна виразити функції (операції) диз'юнкції, кон'юнкції та заперечення. Справді, для будь-якої булевої функції f формулу над Σ, яка зображує f, можна побудувати так. Спочатку слід побудувати булеву формулу для f (за теоремою 6.3 така формула завжди існує) і в цій формулі за допомогою правила заміни всі підформули, що містять операції диз'юнкції, кон'юнкції та заперечення, замінити рівносильними формулами над Σ.

Сформульовану тезу можна узагальнити: якщо всі функції функціонально повної системи Σ' можна виразити за допомогою суперпозицій функцій із системи Σ, то система Σ також функціонально повна. У такому разі будемо говорити, що система Σ *зводиться* до Σ'. Користуючись ***методом зведення***, можна довести функціональну повноту багатьох систем булевих функцій.

Приклад 6.1. 1. Системи $\Sigma_1 = \{\wedge, ^{-}\}$ і $\Sigma_2 = \{\vee, ^{-}\}$ функціонально повні. Справді, за допомогою правил де Моргана і правила подвійного заперечення в кожній із цих систем відсутню порівняно з Σ_0 третю функцію можна виразити через інші дві: $x \vee y = \overline{\overline{x} \wedge \overline{y}}$ і $x \wedge y = \overline{\overline{x} \vee \overline{y}}$.

Наприклад, булева функція, задана булевою формулою $F = xy \vee \overline{x}\,\overline{z} \vee \overline{y}z$ (див. розд. 6.3), над системою Σ_1 дістане зображення $\overline{\overline{xy}\;\overline{\overline{x}\,\overline{z}}\;\overline{\overline{y}z}}$, а над системою Σ_2 — $\overline{\overline{x} \vee \overline{y}} \vee \overline{x \vee z} \vee \overline{y \vee \overline{z}}$.

Це означає, що система твірних $\Sigma_0 = \{\vee\ ,\ \wedge\ ,\ ^{-}\}$ надлишкова й, отже, не є базисом для алгебри $\langle P_2,\ \{S\}\rangle$. Однак, якщо ми відмовимося від надлишковості системи Σ_0, то, як видно з наведеного прикладу, значно ускладнимо вигляд відповідних формул для функцій з P_2. Тому на практиці частіше використовують саме систему Σ_0, ігноруючи її надлишковість.

2. Кожна із систем $\Sigma_3 = \{|\}$ (штрих Шеффера) і $\Sigma_4 = \{\downarrow\}$ (стрілка Пірса) є функціонально повною. Це випливає з таких тотожностей, у справедливості яких неважко переконатися за допомогою стандартного методу:

$$\bar{x} = x \mid x = x \downarrow x,$$
$$xy = \overline{x \mid y} = (x \mid y) \mid (x \mid y),$$
$$x \vee y = \overline{x \downarrow y} = (x \downarrow y) \downarrow (x \downarrow y).$$

Отже, функції системи Σ_1 виражаються через функцію із Σ_3, а функції системи Σ_2 — через функцію із Σ_4, тобто Σ_3 зводиться до Σ_1, а Σ_4 — до Σ_2. Очевидно, що Σ_3 та Σ_4 — базиси для алгебри $\langle P_2,\ \{S\}\rangle$.

3. Особливе місце в алгебрі логіки займає система $\Sigma_5 = \{\ \wedge\ ,\ \oplus\ ,\ 1\}$, яка також функціонально повна. Оскільки $\bar{x} = x \oplus 1$, то система Σ_5 зводиться до Σ_1.

4. Водночас система $\Sigma' = \{\vee\ ,\ \wedge\ ,\ \to\ ,\ \sim\}$, а, отже, і системи $\Sigma'' = \{\vee\ ,\ \wedge\ ,\ \to\}$ і $\Sigma''' = \{\vee\ ,\ \wedge\}$ неповні. У цьому можна переконатися, припустивши супротивне, тобто, що система Σ' функціонально повна. Тоді існує така формула $F(x_1, x_2, \ldots, x_n)$ над Σ', що зображує булеву функцію заперечення $\bar{x}_1$, тобто рівносильна формулі $\bar{x}_1$: $F(x_1, x_2, \ldots, x_n) = \bar{x}_1$. Однак кожна формула над Σ' на наборі значень $(1, 1, \ldots, 1)$ її операндів набуває значення 1, бо кожна з бінарних операцій $\vee\ ,\ \wedge\ ,\ \to$ або $\sim$ набуває значення 1, якщо обидва її операнди дорівнюють 1. Отже, маємо $F(1, 1, \ldots, 1) = 1 \neq \bar{1} = 0$. Дістали суперечність, яка й свідчить про те, що припущення про функціональну повноту системи Σ' неправильне. Схожим методом можна довести неповноту систем $\{\ \vee\ ,\ \oplus\ \}$, $\{\ \wedge\ ,\ \oplus\ \}$, $\{\ \oplus\ ,\ \sim\ \}$ тощо.

Можна продовжувати наводити приклади повних і неповних систем, однак бажано мати ефективний критерій, за яким для довільної системи булевих функцій можна було б визначити, функціонально повна вона чи ні. Інакше кажучи, для алгебри $\langle P_2,\ \{S\}\rangle$ бажано мати *критерій повноти* системи функцій. Такий критерій буде наведено в розд. 12.

6.9. Задачі і вправи

1. Довести, що функцію f, задану формулою A, не можна реалізувати рівносильною формулою над множиною операцій Σ, якщо:

(а) $A = x \oplus y$, $\Sigma = \{\wedge\}$; (в) $A = x \vee y$, $\Sigma = \{\sim\}$;

(б) $A = x \wedge y$, $\Sigma = \{\to\}$; (г) $A = x \wedge y$, $\Sigma = \{\vee\ ,\ \to\}$.

2. Користуючись методом зведення до відомої функціонально повної системи (наприклад, до Σ_0, Σ_1 або Σ_2), довести функціональну повноту системи Σ булевих функцій, заданих такими формулами:

(а) $\Sigma = \{x \to y, \bar{x}\}$;

(б) $\Sigma = \{(x \wedge y) \oplus z, (x \sim y) \oplus z\}$;

(в) $\Sigma = \{x \to y, (\overline{x \oplus y \oplus z})\}$;

(г) $\Sigma = \{x \to y, 0\}$;

(д) $\Sigma = \{x \oplus y, x \vee y, 1\}$;

(е) $\Sigma = \{x \to y, x \oplus y\}$.

3. Довести неповноту системи функцій Σ:

(а) $\Sigma = \{\ \wedge\ ,\ \oplus\ ,\ 0\}$;

(б) $\Sigma = \{\ \wedge\ ,\ \vee\ ,\ \to\ ,\ 1\}$;

(в) $\Sigma = \{\ \wedge\ ,\ \vee\ ,\ \oplus\ ,\ 0\}$.

6.10. Алгебра Жегалкіна

Розглянемо окремо одну важливу систему твірних для алгебри булевих функцій $\langle P_2,\ \{S\}\rangle$, а саме систему $\Sigma_5 = \{\wedge\ ,\ \oplus\ ,\ 1\}$. Її функціональну повноту обґрунтовано у прикладі 6.1.(3) методом зведення.

Поряд із класичною алгеброю логіки $C = \langle B,\ \{\ \wedge\ ,\ \vee\ ,\ ^{-}\}\rangle$, введеною в розд. 4, значне місце в теорії булевих функцій займає алгебра $G = \langle B,\ \{\wedge\ ,\ \oplus\ ,\ 1\}\rangle$ із сигнатурою Σ_5. Вона дістала назву ***алгебри Жегалкіна*** на честь математика І. І. Жегалкіна, який запропонував алгебру G і дослідив її основні властивості.

Як і для алгебри C, подамо ***основні тотожності*** *алгебри Жегалкіна*. Відзначимо тільки властивості операції $\oplus$ та її зв'язок з іншими операціями алгебри G. Усі тотожності з (6.3) щодо операцій $\wedge$ й 1 справедливі, очевидно, і в алгебрі G.

1) $(x \oplus y) \oplus z = x \oplus (y \oplus z)$, *(асоціативність)*

2) $x \oplus y = y \oplus x$, *(комутативність)*

3) $x(y \oplus z) = xy \oplus xz$, *(дистрибутивність)* (6.9)

4) $x \oplus x = 0$, $x \oplus 0 = x$,

5) $x \oplus 1 = \bar{x}$.

Зауважимо, що в алгебрі Жегалкіна введено порядок виконання операцій: спочатку виконуються кон'юнкції, потім — додавання за модулем 2.

Із повноти системи Σ_5 випливає, що будь-яку булеву функцію $f(x_1, x_2, \ldots, x_n)$ можна реалізувати деякою формулою над Σ_5, тобто формулою в алгебрі Жегалкіна. Якщо в цій формулі розкрити всі дужки за законом дистрибутивності, а потім виконати всі спрощення за допомогою основних тотожностей алгебри G, то одержимо рівносильну формулу, яка зображує функцію f і має вигляд

$$\sum_{i_1, i_2, \ldots, i_k} a_{i_1 i_2 \ldots i_k} x_{i_1} x_{i_2} \ldots x_{i_k}, \; a_{i_1 i_2 \ldots i_k} \in B, \; 0 \le k \le n. \tag{6.10}$$

Тут Σ — знак додавання за модулем 2, усі доданки попарно різні, у кожній кон'юнкції жодна змінна не зустрічається більше одного разу. Деякі доданки можуть містити тільки одну змінну ($k = 1$), це лінійні доданки. Нарешті, може бути єдиний доданок із порожньою кон'юнкцією змінних ($k = 0$), яку вважають рівною 1; відповідний коефіцієнт a_0 є вільним членом формули. Рангом порожньої кон'юнкції вважатимемо 0.

Формула (6.10) називається ***поліномом Жегалкіна*** булевої функції $f(x_1, x_2, \ldots, x_n)$. Множники $a_{i_1 i_2 \ldots i_k}$ називають ***коефіцієнтами*** полінома Жегалкіна. Кількість доданків у поліномі Жегалкіна називають його ***довжиною***, а найбільший ранг його елементарної кон'юнкції — ***степенем*** цього полінома.

Теорема 6.7. Будь-яку булеву функцію $f(x_1, x_2, \ldots, x_n)$ можна зобразити у вигляді полінома Жегалкіна, до того ж однозначно.

Доведення. Існування полінома для довільної булевої функції вже доведено. Для доведення єдиності покажемо, що існує взаємно однозначна відповідність між множиною $P_2(n)$ усіх n-арних булевих функцій і множиною всіх поліномів Жегалкіна від n змінних. Кількість різних кон'юнкцій $x_{i_1} x_{i_2} \ldots x_{i_k}$ збігається з кількістю всіх підмножин множини $\{x_1, x_2, \ldots, x_n\}$ і за теоремою 1.1 про потужність булеана скінченної множини дорівнює 2^n (порожній множині відповідає порожня кон'юнкція 1). Кількість різних поліномів, які можна утворити з цих кон'юнкцій, дорівнює кількості всіх підмножин множини кон'юнкцій, і, отже, за тією самою теоремою, вона дорівнює 2^{2^n}. Порожній множині відповідає поліном a_0.

Таким чином, кількість усіх поліномів Жегалкіна від n змінних дорівнює кількості всіх булевих функцій від n змінних. З усіх наведених вище міркувань можна дійти висновку, що відповідність між множиною всіх поліномів Жегалкіна від n змінних і множиною $P_2(n)$ функціональна, всюди визначена та сюр'єктивна. Із рівнопотужності цих множин випливає, що зазначена відповідність ін'єктивна: оскільки кожна функція зображується якимось поліномом, то різні функції мають бути зображувані різними поліномами. Отже, відповідність між множиною $P_2(n)$ усіх n-арних булевих функцій і множиною всіх поліномів Жегалкіна від n змінних взаємно однозначна, що й доводить єдиність полінома Жегалкіна для будь-якої функції.

Поліном Жегалкіна для заданої n-арної булевої функції $f(x_1, x_2, \ldots, x_n)$ можна побудувати так.

1. Побудуємо ДДНФ функції f за формулою (6.7).

2. Оскільки в ДДНФ на кожному з наборів $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, за якими здійснюється диз'юнкція, тільки одна з елементарних кон'юнкцій $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$ дорівнює 1, а решта дорівнюють 0, то у формулі (6.7) операцію диз'юнкції можна замінити операцією додавання за модулем 2.

3. Скориставшись співвідношенням $x^\alpha = x \oplus \alpha$, у справедливості якого легко переконатися, перетворимо одержане в попередньому пункті зображення функції f до вигляду

$$f(x_1, x_2, \ldots, x_n) = \sum_{\substack{(\alpha_1, \alpha_2, \ldots, \alpha_n) \\ f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 1}} (x_1 \oplus \overline{\alpha}_1)(x_2 \oplus \overline{\alpha}_2) \ldots (x_n \oplus \overline{\alpha}_n).$$

4. Розкривши в останній формулі всі дужки за законом дистрибутивності та зробивши всі можливі спрощення за допомогою основних тотожностей алгебри Жегалкіна, дістанемо шуканий поліном.

Приклад 6.2. Побудуємо поліном Жегалкіна для булевої функції, заданої табл. 6.4:

$f(x, y, z) = x^0y^0z^0 \vee x^0y^1z^0 \vee x^1y^0z^0 \vee x^1y^1z^0 \vee x^1y^1z^1 = (x \oplus 1)(y \oplus 1)(z \oplus 1) \oplus$ $\oplus (x \oplus 1)y(z \oplus 1) \oplus x(y \oplus 1)(z \oplus 1) \oplus xy(z \oplus 1) \oplus xyz = (xyz \oplus xz \oplus yz \oplus$ $\oplus z \oplus xy \oplus x \oplus y \oplus 1) \oplus (xyz \oplus yz \oplus xy \oplus y) \oplus (xyz \oplus xy \oplus xz \oplus x) \oplus$ $\oplus (xyz \oplus xy) \oplus xyz = xyz \oplus z \oplus 1$.

Інший спосіб побудови полінома Жегалкіна для булевої функції — це так званий *метод невизначених коефіцієнтів*.

Для заданої булевої функції $f(x_1, x_2, \ldots, x_n)$ запишемо співвідношення

$$f(x_1, x_2, \ldots, x_n) = a_m x_1 x_2 \ldots x_n \oplus a_{m-1} x_1 x_2 \ldots x_{n-1} \oplus a_{m-2} x_1 x_2 \ldots x_{n-2} x_n \oplus \ldots$$
$$\ldots \oplus a_n x_1 \oplus a_{n-1} x_2 \oplus \ldots \oplus a_2 x_{n-1} \oplus a_1 x_n \oplus a_0,$$

права частина якого є поліномом Жегалкіна з невідомими (невизначеними) коефіцієнтами $a_0, a_1, \ldots, a_m$, де $m = 2^n - 1$.

Відтак, підставляючи в обидві частини цього співвідношення всі 2^n наборів α значень булевих змінних $x_1, x_2, \ldots, x_n$ і замінюючи кожне $f(\alpha)$ на відоме значення функції f, отримаємо систему рівнянь відносно $a_i \in B$, $i = 0, 1, 2, \ldots, m$. Розв'язок системи дає шукані коефіцієнти полінома Жегалкіна.

Наприклад, для булевої функції, заданої табл. 6.4, маємо таке співвідношення

$$f(x, y, z) = a_7 xyz \oplus a_6 xy \oplus a_5 xz \oplus a_4 yz \oplus a_3 x \oplus a_2 y \oplus a_1 z \oplus a_0.$$

Тоді

$$f(0, 0, 0) = 1 = a_0,$$
$$f(0, 0, 1) = 0 = a_1 \oplus a_0,$$
$$f(0, 1, 0) = 1 = a_2 \oplus a_0,$$
$$f(0, 1, 1) = 0 = a_4 \oplus a_2 \oplus a_1 \oplus a_0,$$
$$f(1, 0, 0) = 1 = a_3 \oplus a_0,$$
$$f(1, 0, 1) = 0 = a_5 \oplus a_3 \oplus a_1 \oplus a_0,$$
$$f(1, 1, 0) = 1 = a_6 \oplus a_3 \oplus a_2 \oplus a_0,$$
$$f(1, 1, 1) = 1 = a_7 \oplus a_6 \oplus a_5 \oplus a_4 \oplus a_3 \oplus a_2 \oplus a_1 \oplus a_0.$$

Звідси отримуємо $a_0 = a_1 = a_7 = 1$ й $a_2 = a_3 = a_4 = a_5 = a_6 = 0$. Отже, шуканий поліном Жегалкіна — $xyz \oplus z \oplus 1$.

На завершення зауважимо, що поліном Жегалкіна є канонічною формою в класах рівносильних формул. Зведенням будь-яких формул F_1 і F_2 до відповідних поліномів Жегалкіна можна довести чи спростувати твердження про рівносильність цих формул.

6.10. Задачі і вправи

1. Визначити кількість поліномів Жегалкіна степеня k над множиною змінних $X = \{x_1, x_2, \ldots, x_n\}$, $k = 0, 1, 2, \ldots, n$.

2. Визначити кількість поліномів Жегалкіна над множиною змінних $X = \{x_1, x_2, \ldots, x_n\}$, які набувають значення 0 на наборі $(1, 1, \ldots, 1)$.

3. Знайти число поліномів Жегалкіна довжини m над множиною змінних $X = \{x_1, x_2, \ldots, x_n\}$, $m = 1, 2, \ldots, 2^n$.

4. Довести основні тотожності алгебри Жегалкіна.

5. Виходячи з ДДНФ, побудувати поліном Жегалкіна булевої функції $f(x_1, x_2, \ldots, x_n)$, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (1101)$;
(б) $\boldsymbol{a}_f = (01010111)$;
(в) $\boldsymbol{a}_f = (1010111010111011)$.

6. Методом невизначених коефіцієнтів побудувати поліном Жегалкіна булевої функції $f(x_1, x_2, \ldots, x_n)$, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (0111)$;
(б) $\boldsymbol{a}_f = (01010010)$;
(в) $\boldsymbol{a}_f = (1110111110111111)$.

7. Показати, що x_i — істотна змінна функції $f(x_1, x_2, \ldots, x_n)$ тоді й тільки тоді, коли x_i явно входить у поліном Жегалкіна цієї функції.

6.11. Замкнені класи булевих функцій

Система (клас) функцій Ω називається ***замкненою***, якщо суперпозиція будь-яких функцій з Ω належить Ω, тобто $[\Omega]_S = \Omega$. Нагадаємо, що замиканням множини функцій Ω за оператором суперпозиції S (позначають $[\Omega]_S$) називається множина всіх функцій, утворених із функцій множини Ω за допомогою оператора суперпозиції.

Якщо $\Omega \neq P_2$ — замкнений клас булевих функцій і $\Sigma \subseteq \Omega$, то Σ не є функціонально повною системою. Це й визначає те особливе місце, яке посідають у розв'язанні проблеми повноти певні замкнені класи булевих функцій.

Серед замкнених класів булевих функцій особливо важливими є наступні п'ять класів, які іноді називають ***п'ять чудових класів булевих функцій***.

Булева функція $f(x_1, x_2, \ldots, x_n)$ називається функцією, що ***зберігає константу* 0**, якщо $f(0, 0, \ldots, 0) = 0$. Клас усіх булевих функцій, що зберігають константу 0, позначимо T_0.

Булева функція $f(x_1, x_2, \ldots, x_n)$ називається функцією, що ***зберігає константу* 1**, якщо $f(1, 1, \ldots, 1) = 1$. Клас усіх булевих функцій, що зберігають константу 1, позначимо T_1.

Задамо частковий порядок на множині B^n. Для символів двійкового алфавіту B покладемо $0 < 1$. Тоді вважаємо, що двійковий кортеж

$\alpha = (\delta_1, \delta_2, \ldots, \delta_n)$ передує двійковому кортежу $\beta = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ і записуємо $\alpha \le \beta$, якщо $\delta_1 \le \varepsilon_1, \delta_2 \le \varepsilon_2, \ldots, \delta_n \le \varepsilon_n$.

Булева функція $f(x_1, x_2, \ldots, x_n)$ називається ***монотонною***, якщо для будь-яких двійкових кортежів $\alpha, \beta \in B^n$ таких, що $\alpha \le \beta$, виконується нерівність $f(\alpha) \le f(\beta)$. Клас усіх монотонних булевих функцій позначимо через M.

Булева функція $f(x_1, x_2, \ldots, x_n)$, поліном Жегалкіна якої має вигляд $\alpha_0 \oplus (\alpha_1 \wedge x_1) \oplus (\alpha_2 \wedge x_2) \oplus \ldots \oplus (\alpha_n \wedge x_n)$, або, якщо випустити знаки кон'юнкції, — $\alpha_0 \oplus \alpha_1 x_1 \oplus \alpha_2 x_2 \oplus \ldots \oplus \alpha_n x_n$, де $\alpha_i \in B$, $i = 0, 1, 2, \ldots, n$, називається ***лінійною***. Клас усіх лінійних булевих функцій позначимо через L.

Нарешті, нагадаємо, що булева функція $f(x_1, x_2, \ldots, x_n)$ називається ***самодвоїстою***, якщо $f(x_1, x_2, \ldots, x_n) = \overline{f(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n})}$. Клас усіх самодвоїстих булевих функцій позначимо через C.

Приклад 6.3. 1. Функції, що зберігають константу 0, — це, наприклад, диз'юнкція, кон'юнкція, додавання за модулем 2, а заперечення й імплікація — не зберігають константу 0 (див. табл. 6.3).

2. Константу 1 зберігають, наприклад, кон'юнкція, диз'юнкція й імплікація, а функції заперечення та додавання за модулем 2 — не зберігають константу 1.

3. Неважко переконатися, що кон'юнкція та диз'юнкція — монотонні функції, а для заперечення та додавання за модулем 2 умова монотонності не виконується. Наприклад, $(0, 1) \le (1, 1)$, однак $1 \oplus 1 < 0 \oplus 1$.

4. Оскільки $\overline{x} = 1 \oplus x$ і $x \oplus y = 0 \oplus 1x \oplus 1y$, то функції заперечення і додавання за модулем 2 лінійні. Водночас кон'юнкція та диз'юнкція — нелінійні функції. Наприклад, для диз'юнкції маємо $x \vee y = x \oplus y \oplus xy$.

5. Прикладом самодвоїстої функції є функція заперечення.

У наступних лемах доведено замкненість кожного з п'яти означених класів булевих функцій T_0, T_1, M, L і C.

Лема 6.1. Класи T_0 і T_1 булевих функцій, що зберігають константи 0 і 1, замкнені.

Доведення. Для доведення розглянемо будь-яку суперпозицію з T_0 (відповідно з T_1).

Нехай $f_i(x_{i1}, x_{i2}, \ldots, x_{il_i}) \in T_0$, $i = 1, 2, \ldots, k$, $g(z_1, z_2, \ldots, z_k) \in T_0$ і $S(f_1, f_2, \ldots, f_k, g) = g(f_1(x_{11}, x_{12}, \ldots, x_{1l_1}), f_2(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, f_k(x_{k1}, x_{k2}, \ldots, x_{kl_k})) = h(x_1, x_2, \ldots, x_n)$, де $x_1, x_2, \ldots, x_n$ — усі різні символи змінних, які є аргументами функцій $f_1, f_2, \ldots, f_k$. Тоді $h(0, 0, \ldots, 0) = g(f_1(0, 0, \ldots, 0), f_2(0, 0, \ldots, 0), \ldots, f_k(0, 0, \ldots, 0)) = g(0, 0, \ldots, 0) = 0$. Отже, $h \in T_0$.

Аналогічно можна довести замкненість класу T_1.

Лема 6.2. Клас M монотонних булевих функцій замкнений.

Доведення. Нехай $f_i(x_{i1}, x_{i2}, \ldots, x_{il_i}) \in M$, $i = 1, 2, \ldots, k$, $g(z_1, z_2, \ldots, z_k) \in M$. Розглянемо суперпозицію $S(f_1, f_2, \ldots, f_k, g) = g(f_1(x_{11}, x_{12}, \ldots, x_{1l_1}), f_2(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, f_k(x_{k1}, x_{k2}, \ldots, x_{kl_k})) = h(x_1, x_2, \ldots, x_n)$. Як і раніше, $x_1, x_2, \ldots, x_n$ — це всі різні символи змінних, які є аргументами функцій $f_1, f_2, \ldots, f_k$.

Розглянемо два булеві кортежі $\alpha = (\delta_1, \delta_2, \ldots, \delta_n)$ і $\beta = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ довжини n такі, що $\alpha \le \beta$. Позначимо α_i та β_i відповідні проекції кортежів α і β на змінні функції f_i, $i = 1, 2, \ldots, k$. Зрозуміло, що $\alpha_i \le \beta_i$, $i = 1, 2, \ldots, k$. Нехай $\gamma = (f_1(\alpha_1), f_2(\alpha_2), \ldots, f_k(\alpha_k))$ та $\eta = (f_1(\beta_1), f_2(\beta_2), \ldots, f_k(\beta_k))$. Тоді з монотонності функцій $f_1, f_2, \ldots, f_k$ маємо $f_1(\alpha_1) \le f_1(\beta_1), f_2(\alpha_2) \le f_2(\beta_2), \ldots, f_k(\alpha_k) \le f_k(\beta_k)$, тому $\gamma \le \eta$.

Отже, з урахуванням монотонності функції g одержимо $h(\alpha) = h(\delta_1, \delta_2, \ldots, \delta_n) = g(f_1(\alpha_1), f_2(\alpha_2), \ldots, f_k(\alpha_k)) = g(\gamma) \le g(\eta) = g(f_1(\beta_1), f_2(\beta_2), \ldots, f_k(\beta_k)) = h(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = h(\beta)$.

Звідси випливає, що $h \in M$, і замкненість класу M доведено.

Лема 6.3. Клас L лінійних булевих функцій замкнений.

Доведення. Справедливість цієї леми випливає з того, що для будь-яких лінійних булевих функцій $f, g \in L$ і довільної булевої константи $\alpha \in B$ функції $f \oplus g$ і αf також лінійні.

Лема 6.4. Клас C самодвоїстих булевих функцій замкнений.

Доведення. Нехай $f_i(x_{i1}, x_{i2}, \ldots, x_{il_i}) \in C$, $i = 1, 2, \ldots k$, $g(z_1, z_2, \ldots, z_k) \in C$ і суперпозиція $S(f_1, f_2, \ldots, f_k, g) = g(f_1(x_{11}, x_{22}, \ldots, x_{1l_1}), f_2(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, f_k(x_{k1}, x_{k2}, \ldots, x_{kl_k})) = h(x_1, x_2, \ldots, x_n)$. Тоді $h^*(x_1, x_2, \ldots, x_n) = \overline{h(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n})} =$
$= \overline{g}(\overline{f_1(\overline{x_{11}}, \overline{x_{12}}, \ldots, \overline{x_{1l_1}})}, \overline{f_2(\overline{x_{21}}, \overline{x_{22}}, \ldots, \overline{x_{2l_2}})}, \ldots, \overline{f_k(\overline{x_{k1}}, \overline{x_{k2}}, \ldots, \overline{x_{kl_k}})}) =$
$= \overline{g}(\overline{\overline{f_1}(x_{11}, x_{12}, \ldots, x_{1l_1})}, \overline{\overline{f_2}(x_{21}, x_{22}, \ldots, x_{2l_2})}, \ldots, \overline{\overline{f_k}(x_{k1}, x_{k2}, \ldots, x_{kl_k})}) =$
$= g(f_1(x_{11}, x_{22}, \ldots, x_{1l_1}), f_2(x_{21}, x_{22}, \ldots, x_{2l_2}), \ldots, f_k(x_{k1}, x_{k2}, \ldots, x_{kl_k})) =$
$= h(x_1, x_2, \ldots, x_n)$.

Отже, клас C замкнений.

Безпосередньою перевіркою можна укласти таблицю належності елементарних булевих функцій до зазначених п'яти класів (табл. 6.5). Знаком + позначено належність елементарної булевої функції до відповідного класу.

Таблиця 6.5

Назва булевої функції	Позначення	Класи				
		T_0	T_1	M	L	C
Константа 0	0	+	–	+	+	–
Константа 1	1	–	+	+	+	–
Заперечення	$\bar{x}$	–	–	–	+	+
Кон'юнкція	$x \wedge y$	+	+	+	–	–
Диз'юнкція	$x \vee y$	+	+	+	–	–
Додавання за модулем 2	$x \oplus y$	+	–	–	+	–
Еквівалентність	$x \sim y$	–	+	–	+	–
Імплікація	$x \to y$	–	+	–	–	–
Штрих Шеффера	$x \mid y$	–	–	–	–	–
Стрілка Пірса	$x \downarrow y$	–	–	–	–	–

Наявність у кожному стовпчику табл. 6.5 принаймні одного знака "–" свідчить, що жоден із п'яти класів не вичерпує всіх елементарних булевих функцій, а отже, і всіх булевих функцій.

Окрім зазначених п'яти замкнених класів булевих функцій існують й інші замкнені класи в P_2. Це випливає, зокрема, з того, що перетин замкнених класів знову утворює замкнений клас. Однак, як буде показано в наступному розділі, саме класи T_0, T_1, M, L і C мають вирішальне значення для розв'язання проблеми повноти та формулювання критерію повноти для булевих функцій.

6.11. Задачі і вправи

1. Побудувати множину всіх функцій, що залежать від змінних x, y, яка є замиканням множини H:

(а) $H = \{\neg\}$; (в) $H = \{0, \neg\}$; (д) $H = \{f \mid \boldsymbol{a}_f = (01101001)\}$;
(б) $H = \{\oplus\}$; (г) $H = \{\to\}$; (е) $H = \{0, \sim\}$.

2. Довести, що $f \in [H]$, зобразивши f формулою над множиною H:
(а) $f = \neg$, $H = \{0, \to\}$; (б) $f = \oplus$, $H = \{\downarrow\}$; (в) $f = \vee$, $H = \{\to\}$.

3. З'ясувати, які із зазначених нижче множин є замкненими класами:

(а) множина всіх функцій від однієї змінної;

(б) множина всіх таких функцій $f(x_1, x_2, \ldots, x_n)$, що $f(1, 1, \ldots, 1) = 0$;

(в) множина всіх таких функцій $f(x_1, x_2, \ldots, x_n)$, що $|N_f| = 2^{n-1}$;

(г) множина всіх функцій, зображуваних поліномом Жегалкіна не вище другого степеня;

(д) множина всіх функцій, зображуваних у вигляді ДНФ, елементарні кон'юнкції якої не містять заперечень змінних;

(е) множина всіх таких функцій $f(x_1, x_2, \ldots, x_n)$, що $|N_f| = 1$;

(є) множина всіх таких функцій $f(x_1, x_2, \ldots, x_n)$, що $|N_f| = 1$ і $f(1, 1, \ldots, 1) = 0$.

4. Довести, що:

(а) перетин замкнених класів є замкненим класом;

(б) об'єднання двох замкнених класів, узагалі кажучи, не є замкненим класом.

5. Обґрунтувати такі властивості замикання:

(а) $[[K]] = [K]$; (в) $[K_1 \cap K_2] \subseteq [K_1] \cap [K_2]$;
(б) якщо $K_1 \subseteq K_2$, то $[K_1] \subseteq [K_2]$; (г) $[K_1] \cup [K_2] \subseteq [K_1 \cup K_2]$.

6. З'ясувати, якому з класів $T_0 \cup T_1$, $T_1 \setminus T_0$, $T_0 \setminus T_1$ або $T_0 \cap T_1$ належить функція f, задана вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (0100)$;
(б) $\boldsymbol{a}_f = (00000001)$;
(в) $\boldsymbol{a}_f = (0110011000011110)$.

7. З'ясувати, якому з класів $T_0 \cup T_1$, $T_1 \setminus T_0$, $T_0 \setminus T_1$ або $T_0 \cap T_1$ належить функція f, задана формулою A:

(а) $A = ((((x \to (\bar{x} \wedge (z \downarrow y))) \oplus ((y \wedge (x \sim y)) \mid z)) \mid ((x \sim (y \vee z)) \to z)) \vee y)$;

(б) $A = ((((y \downarrow (x \to z)) \mid (((y \to x) \sim \bar{z}) \oplus (x \vee y))) \to x) \mid y)$.

8. З'ясувати, при яких n функція $f(x_1, x_2, \ldots, x_n)$, задана формулою A, належить класу $T_0 \cup T_1$, $T_1 \setminus T_0$, $T_0 \setminus T_1$ або $T_0 \cap T_1$:

(а) $A = x_1 \oplus x_2 \oplus \ldots \oplus x_n$;
(б) $A = x_1x_2 \oplus x_2x_3 \oplus \ldots \oplus x_{n-1}x_n \oplus x_nx_1$;
(в) $A = (x_1 \vee x_2) \oplus (x_1 \vee x_3) \oplus \ldots \oplus (x_1 \vee x_n) \oplus (x_2 \vee x_3) \oplus (x_2 \vee x_4) \oplus \ldots$
$\ldots \oplus (x_{n-1} \vee x_n)$;

(г) $A = \sum x_ix_jx_k$, де суму за модулем 2 обчислено за всіма такими наборами i, j, k, що $1 \le i < j < k \le n$.

9. Довести, що:

(а) $T_0^* = T_1$;

(б) $(T_0 \cup T_1)^* = T_0 \cup T_1$.

10. Підрахувати число функцій, що залежать від змінних $x_1, x_2, \ldots, x_n$ і належать множині:

(а) T_0; (в) $T_0 \cap T_1$; (д) $T_0 \backslash T_1$; (є) $T_0 \Delta T_1$; (з) $P_2 \setminus (T_0 \cap T_1)$;
(б) T_1; (г) $T_0 \cup T_1$; (е) $T_1 \backslash T_0$; (ж) $P_2 \backslash T_0$; (и) $P_2 \setminus (T_0 \cup T_1)$.

11. Зобразивши функцію f, задану вектором значень $\boldsymbol{a}_f$, поліномом Жегалкіна, з'ясувати, чи є вона лінійною:

(а) $\boldsymbol{a}_f = (1011)$;
(б) $\boldsymbol{a}_f = (10101001)$;
(в) $\boldsymbol{a}_f = (0110110000101011)$.

12. З'ясувати, чи є лінійною функція f, задана формулою A:

(а) $A = (x \oplus ((\neg(y \to ((y \vee x) \mid y)) \sim x) \wedge y))$;
(б) $A = ((x \wedge y) \oplus ((x \mid y) \wedge z)) \downarrow ((y \sim z) \to x)$.

13. Замінити у векторі $\boldsymbol{a}$ зірочки символами 0 і 1 так, щоб утворився вектор значень деякої лінійної функції f. Зобразити f поліномом:

(а) $\boldsymbol{a} = (10{*}1)$; (в) $\boldsymbol{a} = (1{*}{*}11{*}0{*})$;
(б) $\boldsymbol{a} = ({*}001{*}1{*}{*})$; (г) $\boldsymbol{a} = (1{*}{*}10{*}1{*})$.

14. Довести, що функція $f(x_1, x_2, \ldots, x_n)$, яка набуває на будь-яких двох сусідніх наборах протилежних значень, лінійна й істотно залежить від усіх своїх змінних. Чи правильне обернене твердження?

15. Довести, що коли функція $f(x_1, x_2, \ldots, x_n)$ лінійна й відмінна від константи, то $|N_f| = 2^{n-1}$. Чи правильне обернене твердження?

16. Знайти число лінійних функцій $f(x_1, x_2, \ldots, x_n)$, що істотно залежать точно від k змінних.

17. Підрахувати число функцій, що залежать від змінних $x_1, x_2, \ldots, x_n$ і належать множині:

(а) L; (г) $T_0 \cup L$; (є) $L \backslash T_1$; (и) $(T_1 \backslash T_0) \cup L$;
(б) $T_0 \cap L$; (д) $T_0 \cap T_1 \cap L$; (ж) $(T_0 \backslash T_1) \cap L$; (і) $P_2 \backslash (T_0 \cup L)$;
(в) $T_1 \cap L$; (е) $T_0 \cup T_1 \cup L$; (з) $L \backslash (T_0 \cup T_1)$; (ї) $P_2 \backslash (T_1 \backslash L)$.

18. Довести, що $L^* = L$.

19. З'ясувати, чи є самодвоїстою функція f, задана вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (0011)$;
(б) $\boldsymbol{a}_f = (00010111)$;
(в) $\boldsymbol{a}_f = (1110001001111000)$.

20. Довести, що вектор значень самодвоїстої булевої функції $f(x_1, x_2, \ldots, x_n)$ має вигляд $(a_0, a_1, a_2, \ldots, a_{m-1}, \overline{a_{m-1}}, \ldots, \overline{a_2}, \overline{a_1}, \overline{a_0})$, де $m = 2^{n-1}$.

21. Довести, що коли $f(x_1, x_2, \ldots, x_n) \in C$, то $|N_f| = 2^{n-1}$. Навести приклад, який показує, що обернене твердження неправильне.

22. З'ясувати, при яких $n \geq 2$ функція $f(x_1, x_2, \ldots, x_n)$, що задається формулою A, самодвоїста:

(а) $A = x_1 \oplus x_2 \oplus \ldots \oplus x_n$;
(б) $A = x_1x_2 \oplus x_2x_3 \oplus \ldots \oplus x_{n-1}x_n \oplus x_nx_1$;
(в) $A = x_1 \oplus x_2 \oplus \ldots \oplus x_n \oplus a$, $a \in B$;
(г) $A = (x_1 \vee x_2) \oplus (x_2 \vee x_3) \oplus \ldots \oplus (x_{n-1} \vee x_n) \oplus (x_n \vee x_1)$.

23. Нехай функції $f(x_1, x_2, \ldots, x_n)$ і $f^*(x_1, x_2, \ldots, x_n)$, $n \geq 1$, задовольняють умову $|N_f| = |N_{f^*}|$. Довести, що:

(а) коли $f \vee f^* \equiv \text{const}$, то $f \in C$;
(б) коли $f \oplus (f \wedge f^*) \equiv \text{const}$, то $f \in C$.

24. Підрахувати кількість самодвоїстих функцій, що залежать від змінних $x_1, x_2, \ldots, x_n$.

25. Підрахувати число функцій, що залежать від змінних $x_1, x_2, \ldots, x_n$ і належать множині:

(а) $T_0 \cap C$; (д) $C \setminus T_0$; (з) $(L \cup C) \setminus T_1$;
(б) $T_1 \cup C$; (е) $C \cap L$; (и) $(C \setminus T_0) \cap T_1$;
(в) $T_0 \cap T_1 \cap C$; (є) $(L \cup T_1) \cap C$; (і) $C \setminus (T_0 \cap T_1 \cap L)$;
(г) $(T_0 \cup T_1) \cap C$; (ж) $L \cup T_0 \cup C$; (ї) $(C \cap L) \setminus (T_0 \Delta T_1)$.

26. Довести, що:

(а) $L \cap T_0 \cap C = L \cap T_1 \cap C = T_0 \cap T_1 \cap L = T_0 \cap T_1 \cap L \cap C$;
(б) $T_0 \cap C = T_1 \cap C = T_0 \cap T_1 \cap C$.

27. Довести, що:

(а) $(T_0 \cup T_1 \cup C \cup L)^* = T_0 \cup T_1 \cup C \cup L$; (в) $(T_0 \setminus C)^* = T_1 \setminus C$;
(б) $((T_0 \cap T_1) \cup C)^* = (T_0 \cap T_1) \cup C$; (г) $(C \setminus T_0)^* = C \setminus T_1$.

28. Довести, що коли функція $f(x_1, x_2, \ldots, x_n)$ немонотонна, то існують два сусідні набори $\boldsymbol{a}$ і $\boldsymbol{b}$ такі, що $\boldsymbol{a} < \boldsymbol{b}$ і $f(\boldsymbol{a}) > f(\boldsymbol{b})$.

29. Довести, що для довільної монотонної функції $f(x_1, x_2, \ldots, x_n)$ справедливі такі формули розкладу:

(а) $f(x_1, x_2, \ldots, x_n) = \left(x_i \wedge f_1^i(x_1, x_2, \ldots, x_n)\right) \vee f_0^i(x_1, x_2, \ldots, x_n)$;
(б) $f(x_1, x_2, \ldots, x_n) = \left(x_i \vee f_0^i(x_1, x_2, \ldots, x_n)\right) \wedge f_1^i(x_1, x_2, \ldots, x_n)$,

де через $f_\alpha^i(x_1, x_2, \ldots, x_n)$ позначено функцію, отриману в результаті підстановки замість змінної x_i значення α.

30. З'ясувати, чи є монотонною функція f, задана вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (0011)$; (в) $\boldsymbol{a}_f = (00110001)$;
(б) $\boldsymbol{a}_f = (00010110)$; (г) $\boldsymbol{a}_f = (0001000101010111)$.

31. Довести, що для довільної монотонної функції f, відмінної від константи, існує ДНФ (КНФ), яка не містить заперечень змінних і реалізує f.

32. Підрахувати число функцій, що залежать від змінних $x_1, x_2, \ldots, x_n$ і належать множині:

(а) $M \setminus (T_0 \cup T_1)$; (в) $M \cap L$; (д) $L \setminus (M \cup C)$.

(б) $M \setminus (T_0 \cap T_1)$; (г) $M \cap C \cap L$;

33. Довести, що $M^* = M$.

34. Довести, що будь-яка монотонна функція міститься не менше ніж у двох із класів T_0, T_1 або L.

35. Довести, що будь-яка функція з L належить принаймні одному з класів T_0, T_1, C або M.

6.12. Теорема про функціональну повноту

Перш ніж формулювати критерій повноти для системи булевих функцій, доведемо кілька допоміжних тверджень.

Лема 6.5. Якщо $f(x_1, x_2, \ldots, x_n)$ — несамодвоїста булева функція, то шляхом підстановки замість її змінних булевих функцій x та $\bar{x}$ тобто за допомогою суперпозиції $S(x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_n}, f)$, можна одержати одну з булевих констант ($\alpha_i \in B$, $i = 1, 2, \ldots, n$).

Доведення. Справді, якщо f — несамодвоїста функція, то існує такий двійковий кортеж $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, що $f(\alpha_1, \alpha_2, \ldots, \alpha_n) = f(\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_n)$. Покладемо $g(x) = S(x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_n}, f) = f(x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_n})$; тоді $g(0) = f(0^{\alpha_1}, 0^{\alpha_2}, \ldots, 0^{\alpha_n}) = f(\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_n) = f(\alpha_1, \alpha_2, \ldots, \alpha_n) = f\left(1^{\alpha_1}, 1^{\alpha_2}, \ldots, 1^{\alpha_n}\right) = g(1)$.

Отже, $g(0) = g(1)$, і функція $g(x)$ — константа.

Лема 6.6. Якщо $\varphi(x_1, x_2, \ldots, x_n)$ — немонотонна булева функція, то шляхом підстановки замість її змінних булевих констант 0 і 1 та булевої функції x можна одержати функцію заперечення $\bar{x}$.

Доведення. Нехай $\varphi(x_1, x_2, \ldots, x_n)$ — немонотонна функція. Тоді існують такі двійкові кортежі $\alpha, \beta \in B^n$, що $\alpha \leq \beta$, але $\varphi(\alpha) > \varphi(\beta)$, тобто $\varphi(\alpha) = 1$ і $\varphi(\beta) = 0$. Якщо кортежі α і β відрізняються k координатами, то в кортежі α відповідні компоненти дорівнюють 0, а в кортежі β — 1.

Утворимо булеву функцію $g(x)$ із функції φ таким чином: усім змінним функції φ, що відповідають однаковим елементам у кортежах α і β, надамо значень цих елементів; замість усіх інших k змінних підставимо булеву функцію x. Тоді $g(0) = \varphi(\alpha) = 1$ і $g(1) = \varphi(\beta) = 0$.

Отже, $g(x) = \bar{x}$.

Лема 6.7. Якщо $\psi(x_1, x_2, \ldots, x_n)$ — нелінійна булева функція, то шляхом підстановки замість її змінних булевих констант можна одержати нелінійну булеву функцію двох аргументів, тобто нелінійну бінарну булеву функцію.

Доведення. Оскільки функція $\psi(x_1, x_2, \ldots, x_n)$ нелінійна, то в її зображенні у вигляді полінома Жегалкіна є доданок $K = x_{j_1} x_{j_2} \ldots x_{j_l}$, який містить кон'юнкцію не менш ніж двох змінних, тобто $l \geq 2$. Покладемо $x_{j_3} = x_{j_4} = \ldots = x_{j_l} = 1$, а всі ті змінні функції ψ, які не входять у K, покладемо рівними 0. Після такої підстановки і зведення подібних членів за законами алгебри Жегалкіна одержимо формулу вигляду $x_{j_1} x_{j_2} \oplus a x_{j_1} \oplus b x_{j_2} \oplus c$, де $a, b, c \in B$, тобто формулу, що задає нелінійну бінарну булеву функцію від двох змінних x_{j_1} і x_{j_2}.

Сформулюємо та доведемо основний результат — *критерій функціональної повноти* системи булевих функцій, який належить відомому математику і логіку Емілю Посту.

Теорема 6.8 (Поста про функціональну повноту). Для того щоб система булевих функцій Σ була функціонально повною, необхідно й достатньо, щоб вона не містилася цілком у жодному з п'яти замкнених класів T_0, T_1, M, L або C.

Інакше кажучи, для того щоб система булевих функцій Σ була функціонально повною, необхідно й достатньо, щоб вона містила:

1) принаймні одну функцію, що не зберігає константу 0;
2) принаймні одну функцію, що не зберігає константу 1;
3) принаймні одну немонотонну функцію;
4) принаймні одну нелінійну функцію;
5) принаймні одну несамодвоїсту функцію.

Доведення. Необхідність. Нехай Σ — функціонально повна система, тоді $[\Sigma]_S = P_2$. Припустімо, що Σ цілком міститься в якомусь із п'яти зазначених класів K, ($\Sigma \subseteq K$). Тоді $P_2 = [\Sigma]_S \subseteq [K]_S = K$. Отже, $K = P_2$. Це неправильно, бо, як зазначено вище, жоден із розглядуваних замкнених класів не збігається з множиною всіх булевих функцій P_2. Необхідність доведено.

Достатність. Нехай система булевих функцій Σ не міститься цілком у жодному з п'яти зазначених класів. Тоді в Σ існує принаймні одна булева функція f_0, що не зберігає константу 0, принаймні одна булева функція f_1, що не зберігає константу 1, принаймні одна немонотонна булева функція f_M, принаймні одна нелінійна булева функція f_l і принаймні одна несамодвоїста булева функція f_C. Враховуючи можливість вилучення і/або введення фіктивних змінних, можна вважати, що всі ці функції f_0, f_1, f_M, f_L і f_C залежать від одних і тих самих змінних $x_1, x_2, \dots, x_n$. Зауважимо також, що ці п'ять функцій не обов'язково різні.

Доведення достатності полягає у зведенні системи Σ до деякої відомої функціонально повної системи булевих функцій: наприклад, до системи $\Sigma_1 = \{ \wedge, ^{-} \}$ (див. приклад 6.1(1)).

Проведемо це доведення в три етапи.

I. Побудуємо за допомогою функцій f_0, f_1 і f_C константи 0 і 1.

Розглянемо функцію f_0. З $f_0 \notin T_0$ випливає, що функція $f_0(x_1, x_2, \dots, x_n)$ не зберігає константу 0, тобто $f_0(0, 0, \dots, 0) = 1$. Далі, залежно від значення функції f_0 на наборі $(1, 1, 1, \dots, 1)$ можливі два випадки.

а). Нехай $f_0(1, 1, \dots, 1) = 1$. Тоді функція $g(x) = f_0(x, x, \dots, x)$ тотожно дорівнює 1, бо $g(0) = f_0(0, 0, \dots, 0) = 1$ і $g(1) = f_0(1, 1, \dots, 1) = 1$.

Іншу константу одержимо за допомогою суперпозиції $S(g, g, \dots, g, f_1) = f_1(g(x), g(x), \dots, g(x)) = h(x)$. Маємо $h(0) = h(1) = f_1(1, 1, \dots, 1) = 0$, оскільки $f_1 \notin T_1$.

б). Нехай $f_0(1, 1, \dots, 1) = 0$. Тоді $\psi(x) = f_0(x, x, \dots, x)$ — це функція заперечення $\bar{x}$, бо $\psi(0) = f_0(0, 0, \dots, 0) = 1$ та $\psi(1) = f_0(1, 1, \dots, 1) = 0$.

Візьмемо $f_C \notin C$. За лемою 6.5 із несамодвоїстої функції f_C та функції $\psi(x) = \bar{x}$ можна одержати одну з функцій-констант $h(x)$. Іншу константу можна утворити за допомогою суперпозиції $S(h, \psi) = \psi(h(x)) = \bar{h}(x)$.

Таким чином, можливість утворення обох нульарних функцій встановлено.

II. Маючи функції-константи 0 і 1 та немонотонну функцію f_M, на підставі леми 6.6 можна одержати функцію заперечення $\psi(x) = \bar{x}$.

III. Нарешті, згідно з лемою 6.7, за допомогою функцій-констант 0 і 1 з нелінійної функції f_L можна побудувати нелінійну функцію вигляду $h(x, y) = xy \oplus ax \oplus by \oplus c$, де коефіцієнти $a, b, c \in B$ залежать від заданої функції f_L.

Зобразимо функцію $h(x, y)$ рівносильними формулами:

$$h(x, y) = (x \oplus b)(y \oplus a) \oplus (ab \oplus c) = x^{\bar{b}} y^{\bar{a}} \oplus (ab \oplus c).$$

Оскільки $(x^{\alpha})^{\alpha} = x$, то, підставивши у функцію h замість x і y відповідно функції $x^{\bar{b}}$, $y^{\bar{a}}$, отримаємо функцію

$$t(x, y) = S(x^{\bar{b}}, y^{\bar{a}}, h) = h(x^{\bar{b}}, y^{\bar{a}}) = (x^{\bar{b}})^{\bar{b}} \oplus (y^{\bar{a}})^{\bar{a}} \oplus (ab \oplus c) = xy \oplus (ab \oplus c).$$

Якщо $ab \oplus c = 0$, то $t(x, y) = xy$. Якщо ж $ab \oplus c = 1$, то $t(x, y) = xy \oplus 1 = \overline{x\,y}$ Тоді, маючи функцію заперечення ψ, побудуємо суперпозицію $S(t, \psi) = \psi(t(x, y)) = \overline{\overline{x\,y}} = xy$.

Отже, за допомогою оператора суперпозиції S із функцій f_0, f_1, f_M, f_L і f_C системи Σ ми побудували булеві функції заперечення $^{-}$ та кон'юнкцію $\wedge$, тобто довели, що система Σ зводиться до функціонально повної системи $\Sigma_1 = \{\wedge, ^{-}\}$. Цим завершується доведення достатності, а отже, й усієї теореми.

Наслідок 6.8.1. Якщо замкнений клас R булевих функцій не збігається з P_2, то він міститься принаймні в одному з класів T_0, T_1, M, L або C.

Доведення теореми про функціональну повноту конструктивне, бо воно містить процедуру (алгоритм) побудови функцій з Σ_1 як суперпозицій функцій з Σ.

Отже, якщо задано якусь функціонально повну систему Σ і довільну булеву функцію $f \in P_2$ $(f \neq 0)$, то формулу над Σ для функції f можна побудувати так. Спочатку побудуємо ДДНФ функції f. За правилами де Моргана виразимо в ній усі операції диз'юнкції через операції кон'юнкції та заперечення. Відтак операції кон'юнкції та заперечення замінимо у формулі суперпозиціями функцій із Σ, користуючись методами доведеної теореми. У результаті дістанемо шукану формулу над Σ.

Таким чином, щоб з'ясувати, чи є дана система Σ булевих функцій функціонально повною, потрібно спочатку побудувати для них таблицю належності до класів T_0, T_1, M, L і C, або так звану ***критеріальну таблицю***, аналогічну табл. 6.5. Система Σ функціонально повна тоді й лише тоді, коли кожен зі стовпчиків одержаної критеріальної таблиці містить принаймні один знак "–". Крім того, функціонально повна система Σ булевих функцій є базисом, якщо після вилучення із системи Σ будь-якої функції, тобто після викреслення відповідного рядка з критеріальної таблиці, з'являється принаймні один стовпчик, який містить лише знаки " + ".

Зокрема, у такий спосіб можна ще раз підтвердити функціональну повноту систем Σ_0, Σ_1, Σ_2, Σ_3, Σ_4 і Σ_5 (приклади 6.1.1–6.1.3), неповноту систем Σ', Σ'' і Σ''' (приклад 6.1.4), переконатися, що кожна із систем

$\{\to, 0\}$, $\{\to, \bar{\ }\}$, $\{\to, \oplus\}$, $\{\sim, \wedge, 0\}$, $\{\sim, \vee, 0\}$ є базисом, а також одержати інші функціонально повні системи та базиси для алгебри булевих функцій $\langle P_2, \{S\}\rangle$.

Система булевих функцій Σ називається ***незалежною***, якщо жодну булеву функцію f із Σ не можна подати формулою над $\Sigma\setminus\{f\}$.

6.12. Задачі і вправи

1. З'ясувати, чи є функціонально повною система Σ булевих функцій, заданих своїми векторами значень:

(а) $\Sigma = \{(0110), (11000011), (10010110)\}$;
(б) $\Sigma = \{(0111), (01011010), (01111111)\}$;
(в) $\Sigma = \{(11), (0111), (00110111)\}$;
(г) $\Sigma = \{(10), (00110111)\}$.

2. З'ясувати, чи є система Σ функцій, що задаються формулами, функціонально повною у P_2:

(а) $\Sigma = \{xy, x \vee y, x \oplus y, xy \vee xz \vee yz\}$;
(б) $\Sigma = \{\bar{x}x(y \sim z) \sim (y \vee z), x \oplus y \oplus z\}$;
(в) $\Sigma = \{x \to y, x \oplus y\}$.

3. Перевірити, чи є базисом у P_2 система Σ функцій, заданих формулами:

(а) $\Sigma = \{x \to y, x \oplus y, x \vee y\}$;
(б) $\Sigma = \{x \oplus y \oplus z, x \vee y, 0, 1\}$;
(в) $\Sigma = \{x \wedge y \vee z, x \wedge y \oplus z, x \wedge y \sim z\}$.

4. Із повної в P_2 системи Σ виділити всі базиси:

(а) $\Sigma = \{0, x \oplus y, x \to y, xy \sim xz\}$;
(б) $\Sigma = \{xy, x \vee y, xy \vee z, x \oplus y, x \to y\}$?

5. Чи є функціонально повною система $\Sigma = \{f, g\}$, якщо:

(а) $f \in C\setminus M$, $g \notin L \cup C$, $f \oplus g \equiv 1$;
(б) $f \notin T_0 \cap T_1$, $g \in M\setminus T_1$, $f \to g \equiv 1$?

6. З'ясувати, чи є функціонально повною система Σ:

(а) $\Sigma = (C \cap M) \cup (L \setminus M)$;
(б) $\Sigma = (L \cap T_1) \cup (C \cap M)$;
(в) $\Sigma = (M \setminus T_0) \cup (L \setminus C)$.

7. Довести, що коли $f \notin T_0 \cup T_1 \cup C$, то $\{f\}$ — функціонально повна система.

8. У множині функцій, які істотно залежать від двох змінних, визначити всі базиси, що містять:

(а) одну функцію; (б) дві функції; (в) три функції.

9. Чи існує у множині функцій, які істотно залежать від двох змінних, базис, що містить чотири функції?

6.13. Алгебра релейно-контактних схем

Однією з основних причин бурхливого розвитку, поширення та зростання значення алгебри логіки в наш час є її застосування в найрізноманітніших галузях науки, техніки та людської діяльності. Зокрема, математичний апарат алгебри логіки широко використовують у сучасній автоматиці й обчислювальній техніці. Це пов'язано з тим, що основними складовими сучасних дискретних автоматичних пристроїв є різноманітні фізичні елементи, які можуть перебувати у двох якісно відмінних стабільних станах. Такі елементи називають ***двопозиційними***, або ***бістабільними***. Один зі станів двопозиційного елемента позначають символом 0, а другий — символом 1.

Найтиповішим двопозиційним елементом є, наприклад, ***реле з одним контактом*** k (рис. 6.2), який може бути замкненим або розімкненим залежно від того, проходить чи не проходить електричний струм у колі AB. У нормальному стані в електричному колі AB немає струму, і контакт k розімкнений. Якщо в колі AB виникає струм, магніт M, намагнічуючись, замикає контакт k.

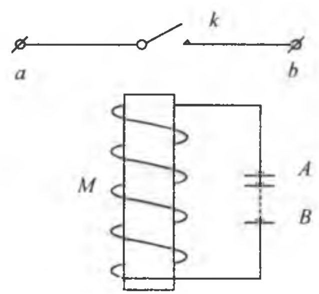


Рис. 6.2

Стану реле "контакт k розімкнений" поставимо у відповідність символ 0, а стану "контакт k замкнений" — символ 1. Оскільки стани "немає струму у колі AB", "магніт M ненамагнічений" і "контакт k розімкнений", а також стани "є струм у колі AB", "магніт M намагнічений" і "контакт k замкнений" рівносильні між собою, то для ідентифікації стану реле достатньо задати один із двох станів контакту k. Тому надалі будемо використовувати спрощене схематичне позначення для реле лише за допомогою контакту (рис. 6.3). Отже, ми ігноруємо зовнішній щодо контакту механізм його замикання та розмикання (ручний або автоматичний перемикач, кнопку вмикання, обмотку реле тощо) і залишаємо й аналізуємо тільки сам контакт і його стан.

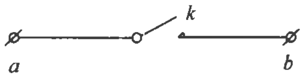


Рис. 6.3

Розглядуване реле і відповідний контакт називають ***реле замикання*** та ***контакт замикання***. Використовують також так звані ***реле розмикання*** і відповідний ***контакт розмикання***. У нормальному положенні такий контакт замкнений. Якщо в колі AB з'являється струм і магніт M намагнічується, контакт розмикається. Схематично контакт розмикання будемо позначати як на рис. 6.4.

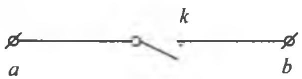


Рис. 6.4

Оскільки кожен контакт k може перебувати у двох станах 0 або 1, то йому можна поставити у відповідність деяку булеву змінну x. Два чи більше різних контактів, які одночасно замикаються чи розмикаються під дією одного й того самого зовнішнього механізму (тобто контакти, які завжди перебувають в однакових станах), позначатимемо одним символом. Цим контактам відповідає одна й та сама булева змінна. Іноді на схемах такі контакти з'єднують між собою пунктирними лініями.

Окремо виділимо два особливі контакти: *постійно замкнений* і *постійно розімкнений*. Цим контактам відповідатимуть булеві константи 1 і 0. Схематичне їх позначення зображено на рис. 6.5.

Рис. 6.5

Реле замикання, зображене на рис. 6.2, використовують як комутаційний пристрій для встановлення (контакт k замкнений) або розриву (контакт k розімкнений) зв'язку між точками a і b якогось електричного, інформаційного, технологічного тощо ланцюга. Ті самі задачі виконує й реле розмикання. Відсутність або наявність зв'язку між точками a і b, залежно від стану контакту k, будемо називати ***результатом дії***, або ***результатом вихідної дії***, реле та позначатимемо відповідно через 0 і 1 відповідно.

Розглянемо на множині контактів (замикання і розмикання) такі *операції*.

1. ***Додавання контактів***. Сумою $k_1 + k_2$ контактів k_1 і k_2, яким відповідають булеві змінні x_1 та x_2, називатимемо паралельне сполучення цих контактів, як зображено на рис. 6.6.

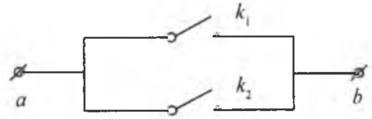


Рис. 6.6

Узагальнюючи поняття результату дії реле, вважатимемо результатом дії одержаної схеми відсутність або наявність зв'язку між точками a і b. Неважко переконатися, що результат дії додавання контактів дорівнює 1 тоді й тільки тоді, коли принаймні один із контактів операндів k_1 або k_2 замкнений, тобто коли результат дії принаймні одного з реле дорівнює 1. Отже, результат дії додавання контактів дорівнює диз'юнкції відповідних булевих змінних — $x_1 \vee x_2$.

2. ***Множення контактів***. Добутком $k_1 * k_2$ контактів k_1 і k_2, яким відповідають булеві змінні x_1 і x_2, називатимемо послідовне сполучення цих контактів, яке зображено на рис. 6.7.

Рис. 6.7

Результат дії множення контактів дорівнює кон'юнкції відповідних булевих змінних — $x_1 \wedge x_2$.

3. ***Запереченням*** $\neg k$ контакту k, якому відповідає булева змінна x, є контакт розмикання, якщо контакт k — контакт замикання,

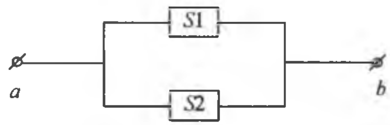
і навпаки. Результат дії заперечення контакту k збігається із запереченням булевої змінної — $\bar{x}$.

Об'єкти, одержані в результаті виконання операцій додавання, множення і заперечення над контактами замикання та розмикання, називатимемо ***релейно-контактними (РК) схемами***. Як у довільних алгебрах окремі змінні можна вважати найпростішими виразами, так і окремі контакти вважатимемо тривіальними релейно-контактними схемами і називатимемо ***елементарними релейно-контактними схемами***.

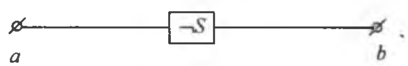
Введені операції додавання, множення і заперечення можна узагальнити. Разом з узагальненням операцій індуктивно вводять поняття релейно-контактної схеми, а також поняття результату дії та булевої формули, яка відповідає довільній релейно-контактній схемі.

1. Вважаємо релейно-контактними схемами всі елементарні релейно-контактні схеми.

2. Якщо $S1$ і $S2$ деякі релейно-контактні схеми, то їх сумою $S1 + S2$ є схема а добутком $S1 * S2$ — схема



Запереченням релейно-контактної схеми S є схема, результатом дії якої є 1 тоді й тільки тоді, коли результатом дії схеми S є 0, і навпаки. Заперечення схеми S позначимо $\neg S$:



Можна сформулювати такий алгоритм побудови релейно-контактної схеми $\neg S$, що є запереченням даної релейно-контактної схеми S. Для цього у схемі S усі контакти замикання треба замінити на контакти розмикання, і навпаки. Відтак усі паралельні з'єднання замінити на відповідні послідовні, й навпаки.

Отже, можна розглядати алгебру $R = \langle C, \{ + , *, \neg\}\rangle$, носієм C якої є множина всіх релейно-контактних схем. Алгебру R називають ***алгеброю релейно-контактних схем***.

Вважатимемо, що ***результат дії*** релейно-контактної схеми S дорівнює 1, якщо схема S встановлює зв'язок між точками a і b, і дорівнює 0 в іншому випадку.

Індуктивне правило побудови релейно-контактних схем задає також алгоритм, який кожній релейно-контактній схемі зіставляє деяку булеву формулу.

1. Кожному контакту поставимо у відповідність деяку булеву змінну (елементарну формулу).

2. Нехай релейно-контактним схемам $S1$ і $S2$ відповідають булеві формули $F1$ і $F2$. Результат дії схеми $S1 + S2$ дорівнюватиме 1 тоді й тільки тоді, коли дорівнює 1 результат дії принаймні однієї зі схем $S1$ або $S2$. Отже, результат дії релейно-контактної схеми $S1 + S2$ задає булева формула $F1 \vee F2$.

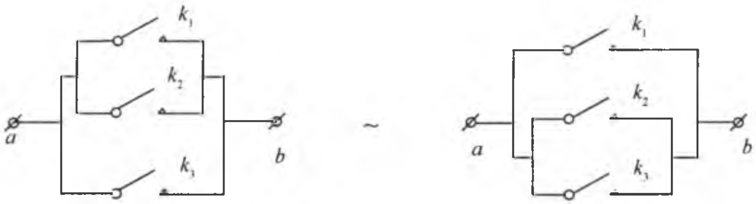
Аналогічно можна виявити, що результат дії релейно-контактної схеми $S1 * S2$ задає булева формула $F1 \wedge F2$. Результат дії схеми $\neg S1$ задає булева формула $F1'$, яку отримуємо з формули $\overline{F1}$, застосовуючи до неї правила де Моргана і правило подвійного заперечення доти, доки заперечення не стоятимуть над окремими змінними.

Позначимо через γ описане вище відображення множини C релейно-контактних схем у множину Φ_B булевих формул, за яким кожній релейно-контактній схемі S ставиться у відповідність булева формула F, що задає результат дії схеми S. Із наведених вище міркувань випливає, що γ — ізоморфне відображення алгебри релейно-контактних схем $R = \langle C, \{+ , *, \neg\}\rangle$ в алгебру булевих формул $D'_B = \langle \Phi'_B, \{\vee , \wedge , ^-\}\rangle$, де Φ'_B — це множина булевих формул, у яких заперечення стоять лише над окремими змінними.

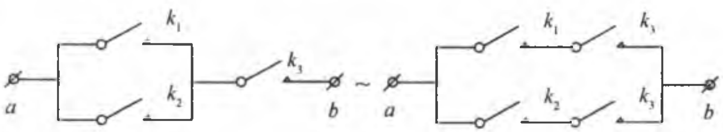
Дві релейно-контактні схеми $S1$ і $S2$ назвемо ***еквівалентними***, якщо для будь-яких варіантів замикання і розмикання відповідних їх контактів результати дії обох схем збігаються. Еквівалентність схем позначатимемо знаком $\sim$: $S1 \sim S2$. Неважко переконатися, що релейно-контактні схеми еквівалентні тоді й тільки тоді, коли відповідні їм булеві формули рівносильні.

З означення еквівалентності релейно-контактних схем та існування ізоморфізму між алгебрами R і D'_B випливає низка тверджень стосовно алгебри релейно-контактних схем R. По-перше, алгебра R — булева. По-друге, для її операцій виконуються тотожності, аналогічні основним тотожностям алгебри логіки (6.3). Нарешті, в алгебрі R можна застосовувати рівносильні перетворення, за допомогою яких з одних релейно-контактних схем можна одержувати інші еквівалентні їм релейно-контактні схеми. Наприклад, за допомогою рівносильних перетворень можна перевіряти еквівалентність релейно-контактних схем.

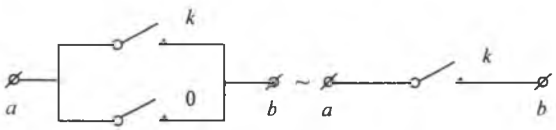
Деякі пари еквівалентних релейно-контактних схем подано на рис. 6.8. Під кожною з них виписано відповідну рівносильність для булевих формул.



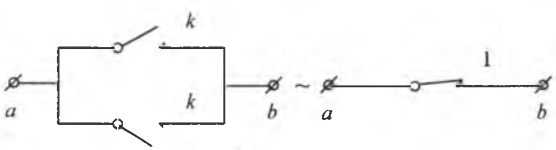
$(x_1 \vee x_2) \vee x_3 = x_1 \vee (x_2 \vee x_3)$



$(x_1 \vee x_2) \wedge x_3 = (x_1 \wedge x_3) \vee (x_2 \wedge x_3)$



$x \vee 0 = x$



$x \vee \bar{x} = 1$

Рис. 6.8

Зрозуміло, що всі рівносильні перетворення в алгебрі релейно-контактних схем, зокрема перевірку їх еквівалентності, зручніше виконувати не безпосередньо над схемами, а опосередковано — за допомогою відповідних їм булевих формул.

Кожній релейно-контактній схемі можна поставити у відповідність деяку булеву функцію. Це можна зробити за допомогою відображення γ: $C \to \Phi'_B$. Релейно-контактній схемі S ставимо у відповідність булеву функцію f, яку реалізує відповідна формула $F = \gamma(S)$. Однак часто корисно мати можливість будувати булеву функцію f безпосередньо за заданою схемою S (не використовуючи відповідної булевої формули).

Нехай задано релейно-контактну схему S. Поставимо у відповідність усім різним контактам k_1, k_2, …, k_n схеми S булеві змінні x_1, x_2, …, x_n. Якщо схема S містить n різних контактів, то існує 2^n різних варіантів замикання і розмикання цих контактів: від ситуації, коли всі контакти розімкнено, до варіанта, коли всі n контактів замкнено. Кожний із цих варіантів визначає певний результат дії схеми S: 0 або 1. Кожному з цих варіантів відповідає також набір значень булевих змінних x_1, x_2, …, x_n: від набору (0, 0, …, 0) до набору (1, 1, …, 1). Виписавши всі ці набори та відповідні їм результати дії схеми, одержимо таблицю істинності шуканої булевої функції $f(x_1, x_2, …, x_n)$.

Побудована так для заданої релейно-контактної схеми S булева функція f — це функція, яку реалізує відповідна булева формула $F = \gamma(S)$. Отже, еквівалентним релейно-контактним схемам відповідає одна й та сама булева функція. Це означає, що еквівалентні схеми "поводяться" однаково, тобто мають однакові результати дії. Враховуючи це, часто вважають, що для задання релейно-контактної схеми (точніше, для задання умов її роботи чи алгоритму дії) достатньо задати відповідну булеву функцію.

Відзначимо ***основні проблеми*** алгебри релейно-контактних схем.

1. ***Аналіз*** схеми. Для заданої релейно-контактної схеми визначити алгоритм її дії, тобто побудувати відповідну булеву функцію.

2. ***Синтез*** схеми. Побудувати релейно-контактну схему за заданими умовами її роботи (наприклад, за заданою булевою функцією чи відповідною булевою формулою).

3. ***Оптимізація***. Розв'язок задачі синтезу не єдиний: існує безліч релейно-контактних схем, які відповідають заданим умовам роботи. За-

дача оптимізації полягає в тому, щоб серед усіх еквівалентних між собою релейно-контактних схем, для яких виконуються задані умови роботи, знайти ту, яка найкраще задовольняє певні вимоги. Найчастіше вимагають, щоб схема містила найменшу можливу кількість елементів (контактів).

Задача аналізу найпростіша з усіх трьох. Метод її розв'язання — описана вище процедура побудови таблиці істинності шуканої булевої функції безпосередньо за заданою релейно-контактною схемою. Другий метод полягає в послідовній побудові за заданою релейно-контактною схемою відповідної їй булевої формули, а за одержаною формулою — таблиці істинності шуканої булевої функції.

Перш ніж формулювати алгоритм розв'язання ***задачі синтезу***, потрібно з'ясувати таке питання: чи для будь-яких заданих умов роботи (чи для довільної булевої функції) існує релейно-контактна схема, яка задовольняє ці умови?

Відповідь на це питання випливає з двох відомих фактів: по-перше, з існування для довільної булевої функції відповідної їй булевої формули спеціального вигляду і, по-друге, з ізоморфізму між алгеброю булевих формул D'_B і алгеброю релейно-контактних схем R. З ізоморфізму між D'_B і R можна зробити висновок, що система операцій $\{+ , *, \neg\}$ функціонально повна для алгебри релейно-контактних схем. Для будь-якої булевої функції можна побудувати відповідну їй релейно-контактну схему з елементарних схем за допомогою паралельних і/або послідовних з'єднань та операції перетворювання даної схеми S на схему $\neg S$.

Підсумовуючи вищезазначене, сформулюємо *алгоритм синтезу* релейно-контактної схеми за заданими умовами її роботи у вигляді деякої булевої функції f.

1. За функцією f побудувати булеву формулу F, що реалізує f. За допомогою правил де Моргана і правила подвійного заперечення перетворити формулу F у рівносильну їй формулу F', у якій заперечення стоятимуть тільки над окремими змінними.

2. Булеву формулу F' розкласти на підформули, тобто подати у вигляді $F_1 \vee F_2$ або $F_1 \wedge F_2$. У першому випадку побудувати схему $S_1 + S_2$, у другому — схему $S_1 * S_2$. Тут S_1 і S_2 — схеми, що відповідають формулам F_1 і F_2.

3. Кожну з підформул F_1 і F_2, у свою чергу, розкласти на підформули, а схеми S_1 і S_2 подати у вигляді з'єднання (композиції) відповідних підсхем.

4. Продовжувати цей процес доти, доки всі підформули не можна буде розкласти, а відповідні їм підсхеми не стануть елементарними релейно-контактними схемами.

Побудована в такий спосіб схема є шуканою релейно-контактною схемою, розв'язком задачі синтезу для булевої функції f.

Ще один варіант методу синтезу релейно-контактних схем можна дістати, використовуючи ізоморфізм між алгебрами D'_B і R.

Застосовуючи попередній алгоритм синтезу паралельно з побудовою схеми, що відповідає булевій функції f, у графічному вигляді, одержимо вираз W в алгебрі релейно-контактних схем R, який можна вважати аналітичним записом цієї схеми. Цей вираз можна одержати й безпосередньо з булевої формули F', виходячи зі згаданого ізоморфізму. Для цього кожній змінній x_i формули F' поставимо у відповідність контакт k_i: контакт размикання тим змінним, які входять до формули F' без заперечення та контакт замикання — змінним, які входять в F' із запереченням. Відтак, операції диз'юнкції у формулі F' відповідатиме операція додавання у виразі W, а операції кон'юнкції — операція множення. Тепер графічний варіант шуканої схеми можна відтворити за виразом W.

Аналізуючи наведені алгоритми синтезу, можна дійти висновку, що *складність* синтезованої релейно-контактної схеми, тобто кількість контактів у ній, дорівнює кількості входжень змінних, або кількості операндів у формулі F' або у відповідному виразі W. Отже, задача оптимізації релейно-контактної схеми за критерієм мінімальності використовуваних контактів зводиться до задачі побудови для заданої булевої функції f такої булевої формули F, яка містить найменшу можливу кількість операндів порівняно з усіма іншими булевими формулами, що реалізують булеву функцію f.

Розглянемо простий приклад, який ілюструє сформульовані методи і висновки.

Приклад 6.4. Комітет складається з трьох осіб і має намір застосовувати релейно-контактну схему для реалізації таємного голосування за принципом більшості голосів. Передбачено, що кожний член комітету, який голосує "за", натискає кнопку на своєму столі. Якщо більшість членів комітету проголосує "за", має засвітитися сигнальна лампочка.

Позначимо через k_1, k_2 і k_3 контакти кнопок, що належать трьом членам комітету, а через x_1, x_2 і x_3 — відповідні булеві змінні. Результат дії шуканої релейно-контактної схеми дорівнює 1 тоді й тільки тоді, коли значення принаймні двох булевих змінних дорівнює 1. Отже, таблиця істинності булевої функції f, яка описує алгоритм дії шуканої схеми, має такий вигляд (табл. 6.6).

Таблиця 6.6

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Їй відповідають булева формула

$$F = \overline{x}_1 x_2 x_3 \vee x_1 \overline{x}_2 x_3 \vee x_1 x_2 \overline{x}_3 \vee x_1 x_2 x_3$$

і вираз

$$W = ((\neg k_1) * k_2 * k_3) + (k_1 * (\neg k_2) * k_3) + (k_1 * k_2 * (\neg k_3)) + (k_1 * k_2 * k_3).$$

Застосувавши один зі сформульованих алгоритмів синтезу, одержимо релейно-контактну схему, зображену на рис. 6.9.

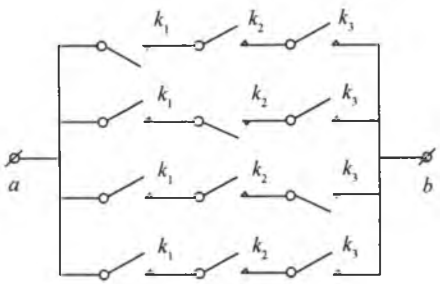


Рис. 6.9

Одержану схему можна спростити (оптимізувати) за допомогою рівносильних перетворень формули F. Скористаємося властивістю ідемпотентності диз'юнкції і двічі продублюємо у формулі F доданок $x_1 x_2 x_3$. Одержимо

$$F = \overline{x}_1 x_2 x_3 \vee x_1 \overline{x}_2 x_3 \vee x_1 x_2 \overline{x}_3 \vee x_1 x_2 x_3 = (\overline{x}_1 x_2 x_3 \vee x_1 x_2 x_3) \vee$$
$$\vee (x_1 \overline{x}_2 x_3 \vee x_1 x_2 x_3) \vee (x_1 x_2 \overline{x}_3 \vee x_1 x_2 x_3) = x_2 x_3 (\overline{x}_1 \vee x_1) \vee x_1 x_3 (\overline{x}_2 \vee x_2) \vee$$
$$\vee x_1 x_2 (\overline{x}_3 \vee x_3) = x_2 x_3 \cdot 1 \vee x_1 x_3 \cdot 1 \vee x_1 x_2 \cdot 1 = x_1 x_2 \vee x_1 x_3 \vee x_2 x_3.$$

Відповідний вираз в алгебрі схем — $W = (k_1 * k_2) + (k_1 * k_3) + (k_2 * k_3)$.

Отже, сформульовані умови роботи задовольняє також релейно-контактна схема, зображена на рис. 6.10.

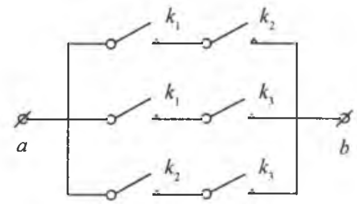


Рис. 6.10

Таким чином, за допомогою простих рівносильних перетворень вихідної булевої формули нам вдалося скоротити вдвічі (з 12 до 6) кількість контактів у синтезованій схемі. До речі, останню схему можна перетворити в еквівалентну релейно-контактну схему, що містить лише п'ять контактів.

Разом зі зменшенням об'єму та вартості спрощення (оптимізація) схеми за кількістю контактів спричиняє також підвищення надійності й ефективності роботи релейно-контактної схеми.

Ідею застосування алгебри логіки для дослідження релейно-контактних схем висловив ще в 1910 р. відомий фізик П. Еренфест. Однак тривалий час ця ідея лишалася нереалізованою. Для проектування, аналізу й оптимізації простих схем інженерам вистачало повсякденного досвіду й інтуїції, і вони не мали потреби у застосуванні загальної теорії і математичного апарату. Ситуація істотно змінилася в 30-х роках XX століття. На той час автоматичні пристрої, створювані на базі релейно-контактних схем, досягли такого рівня складності, що інтуїтивні та "ручні" методи їх побудови й аналізу вже не могли забезпечити бажаного рівня ефективності та якості. З тих пір застосування математичного апарату алгебри логіки в теорії та практиці автоматичних пристроїв постійно розширюється. Ця тенденція особливо підсилилася в епоху створення ЕОМ і виникнення кібернетики [7, 8, 14].

6.13. Задачі і вправи

1. Побудувати релейно-контактну схему (РК схему), що відповідає формулі A алгебри релейно-контактних схем:

(а) $A = ((a + \neg b) * a + (\neg b + \neg a)) * \neg a$;

(б) $A = (a + (a + \neg b) * (((b + c) + \neg a) * (\neg c + a) + b) * \neg a)$.

2. За заданою формулою A алгебри релейно-контактних схем визначити вектор значень булевої функції f, що описує результат дії відповідної релейно-контактної схеми:

(а) $A = (((a * (\neg b + \neg a)) + ((a + \neg b) * a) + a) + \neg b)$;

(б) $A = (a + \neg b + \neg c) * (a + (\neg c * (c + b * \neg a) + b))$.

3. Накреслити РК схему, результат дії якої задано вектором значень $\boldsymbol{a}_f$ булевої функції f:

(а) $\boldsymbol{a}_f = (0010)$;

(б) $\boldsymbol{a}_f = (00011001)$;

(в) $\boldsymbol{a}_f = (1001011100001100)$.

4. Накреслити РК схему, що складається з k контактів і результат дії якої задає формула A алгебри релейно-контактних схем:

(а) $A = a * b * c + a * \neg b * c + \neg a * b * c$, $k = 3$;

(б) $A = a * b * \neg c + \neg a * b + b * \neg c + \neg b$, $k = 3$;

(в) $A = (b + c) * (\neg a + b) * (a + c)$, $k = 4$;

(г) $A = (b + c) * (a + b) * (\neg a + c)$, $k = 4$.

5. Побудувати РК схему, що автоматизує процедуру голосування комітету, який складається з трьох членів, один з яких — голова. Усі рішення приймають більшістю голосів за єдиним винятком: рішення вважають неприйнятим, якщо голова комітету не проголосував "за".

6. Побудувати РК схему з контактами k_1, k_2, k_3 і k_4, результат дії якої є 1, коли замкнено:

(а) будь-який з контактів; (в) усі чотири контакти;

(б) будь-які два контакти; (г) більш як два контакти.

7. Екзаменаційний білет містить чотири питання, на які потрібно відповідати "так" або "ні". Студент складає залік, якщо на перші три питання він відповідає "так" або якщо на друге та четверте питання він відповідає "ні". Побудувати РК схему, яка прийматиме залік, причому відповіді "так" відповідає замикання відповідного контакту.

6.14. Алгебра комбінаційних схем

Як уже неодноразово зазначалося вище, у сучасній техніці, зокрема в автоматиці й обчислювальній техніці, надзвичайно поширені ***двопозиційні (бістабільні) елементи***. Один із двох стабільних станів таких елементів позначають символом 0, а другий — символом 1. Отже, можна вважати, що кожний двопозиційний елемент реалізує булеву змінну. На сьогодні існує досить багато різних способів фізичної реалізації двопозиційних елементів, тобто інтерпретації булевих змінних. Прикладами такої інтерпретації можуть бути: високий потенціал (напруга) — нульовий потенціал, є імпульс — немає імпульсу, позитивний імпульс — негативний імпульс, протилежні за знаком значення магнітної індукції тощо.

Для кожного способу фізичної інтерпретації булевих змінних використовують відповідні набори елементів, які реалізують основні операції алгебри логіки над цими змінними. Такі елементи називають ***комбінаційними***, або ***логічними***.

Принципи побудови і функціонування та технологія виготовлення кобінаційних елементів постійно розвиваються та вдосконалюються. Комбінаційні елементи перших ЕОМ було створено на базі електронних ламп. Сучасні комбінаційні елементи являють собою мікромініатюризовані інтегральні електронні схеми (мікросхеми), сформовані в кристалі кремнію за допомогою спеціальних технологічних процесів.

Розглядаючи комбінаційні елементи, ми залишаємо осторонь принцип їх дії та фізичні процеси, що проходять в них. Для нас достатньо того, що існують якісь фізичні величини, котрі інтерпретують булеві змінні. Вони подаються на вхід комбінаційного елемента; при цьому на виході елемента з'являється певна фізична величина, яку теж можна інтерпретувати як булеву змінну.

На схемах комбінаційні елементи позначають прямокутниками, у середині яких вписують символи чи назви операцій, що реалізують ці елементи. Стрілками чи лініями позначають входи і виходи комбінаційних елементів. Біля цих стрілок можуть бути записані символи, якими позначено відповідні булеві змінні. Приклади комбінаційних елементів, що реалізують основні операції алгебри логіки, зображено на рис. 6.11.

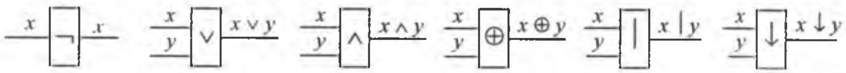


Рис. 6.11

У кожній конкретній галузі техніки використовують більш або менш стандартизовані набори комбінаційних елементів. Позначимо через E множину використовуваних комбінаційних елементів. Булеву функцію, яку реалізує комбінаційний елемент $e \in E$, позначатимемо φ_e.

Маючи комбінаційні елементи, що реалізують певні булеві функції, можна будувати схеми для реалізації складніших булевих функцій.

Інструментом побудови довільних булевих функцій з деяких вихідних булевих функцій є оператор суперпозиції S. Неважко встановити технічний аналог оператора суперпозиції, за допомогою якого з комбінаційних елементів можна одержувати складніші схеми. А саме, з'єднання комбінаційних елементів, при якому вихід одного елемента з'єднано з входом іншого, відповідає підстановці в булеву функцію замість аргумента іншої булевої функції, тобто оператору суперпозиції. Перекомутація входів комбінаційного елемента відповідає перестановці аргументів булевої функції. Об'єднанню деяких входів комбінаційних елементів відповідає ототожнення відповідних булевих змінних.

За аналогією з оператором суперпозиції S для множин функцій неважко індуктивно означити операцію з'єднання Z на множині комбінаційних елементів E. Результати застосування операції з'єднання Z на множині E будемо називати ***комбінаційними схемами*** над E. Множину всіх комбінаційних схем над E позначимо K_E. Отже, $K_E = [E]_Z$. Алгебру $L_E = \langle K_E, \{Z\} \rangle$ будемо називати ***алгеброю комбінаційних схем***.

Нехай k — деяка комбінаційна схема. Назвемо вільними входами комбінаційних елементів ті входи, до яких не приєднано жодного з виходів комбінаційних елементів схеми. Усі вільні входи комбінаційних елементів утворюють входи комбінаційної схеми k. Єдиний вільний вихід є виходом усієї схеми k.

Кожна комбінаційна схема k є фізичною реалізацією якоїсь булевої функції φ_k. Для побудови шуканої функції позначимо всі входи схеми k символами булевих змінних. Аналітичний вираз (функціональний або алгебричний) для φ_k одержимо, послідовно інтерпретуючи кожне зі з'єднань комбінаційних елементів суперпозицією відповідних булевих функцій.

Описана процедура побудови булевої функції, яку реалізує задана комбінаційна схема, є процедурою розв'язання так званої ***задачі аналізу*** схеми, тобто задачі опису алгоритму дії заданої схеми. Наприклад, для комбінаційної схеми, зображеної на рис. 6.12, функціональний і алгебричний вирази для булевої функції, яку реалізує ця схема, мають відповідно вигляд $\varphi_{13}(\varphi_1(\psi_2(\varphi_7(\psi_2(x), y)), z), \ \varphi_6(x, y))$ і $(\overline{\bar{x} \vee y}) \wedge z \rightarrow (x \oplus y)$.

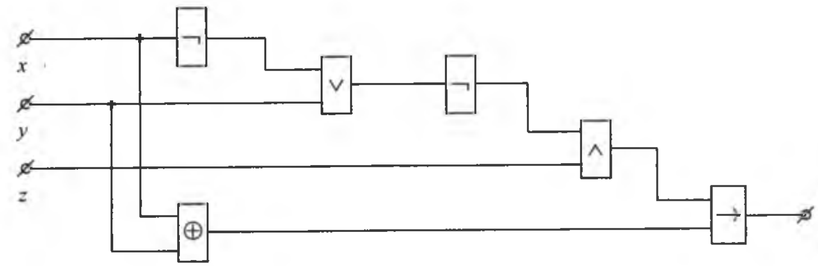


Рис. 6.12

Протилежною є ***задача синтезу***, тобто задача побудови комбінаційної схеми, що реалізує задані умови роботи. Умови роботи комбінаційної схеми задають зазвичай у вигляді певної булевої функції f, яку має реалізувати шукана схема. Перша проблема, що постає при розв'язанні задачі синтезу, — це проблема існування її розв'язку. Чи завжди для даної системи комбінаційних елементів E можна побудувати комбінаційну схему над E, яка реалізує задану булеву функцію f? Очевидно, позитивну відповідь на це запитання одержимо тільки тоді, коли системі E відповідає функціонально повна система булевих функцій. Природно назвати таку систему комбінаційних елементів ***повною***.

Якщо E — якась система комбінаційних елементів, то алгебра комбінаційних схем $\langle K_E, \{Z\} \rangle$ ізоморфна алгебрі булевих функцій $\langle \Phi_E, \{S\} \rangle$, де $\Phi_E = \{\varphi_e \mid e \in E\}$. Нехай E — повна система. Виходячи із зазначеного ізоморфізму, алгоритм синтезу комбінаційної схеми k, яка реалізує задану булеву функцію f, можна сформулювати так. Використавши доведення теореми 6.8 про функціональну повноту, потрібно знайти формулу F (функціональну чи алгебричну) над функціонально повною системою булевих функцій Φ_E, що зображує булеву функцію f. Відтак за формулою F слід побудувати шукану комбінаційну схему k.

Нарешті, третьою основною задачею алгебри комбінаційних схем, істотно складнішою від задач аналізу і синтезу, є задача оптимізації комбінаційної схеми. Комбінаційні схеми k_1 і k_2 назвемо ***еквівалентними***, якщо вони реалізують ту саму булеву функцію. ***Задача оптимізації***, або задача оптимального синтезу, для комбінаційних схем полягає в тому, щоб за заданими умовами роботи побудувати відповідну комбінаційну схему, оптимальну за певними критеріями. Зазвичай таким критерієм обирають складність схеми, під якою розуміють кількість комбінаційних елементів, що входять у неї. Зрозуміло, що розв'язання задачі оптимального синтезу зводиться до пошуку для заданої булевої функції f, що описує умови роботи шуканої схеми, формули F над множиною Φ_E, яка зображує функцію f і містить найменшу можливу кількість операцій (елементів φ_e, $e \in E$).

На практиці часто повні системи комбінаційних елементів є надлишковими. Це дає змогу будувати схеми, оптимальні за кількістю використовуваних елементів.

Через E_0 позначатимемо множину комбінаційних елементів, що відповідають операціям $\vee$, $\wedge$, $\neg$. Часто, при визначенні складності комбінаційної схеми над E_0, елементи, що відповідають операції заперечення, не враховують.

6.14. Задачі і вправи

1. Побудувати комбінаційну схему, що реалізує булеву функцію f, задану вектором значень $\boldsymbol{a}_f$. Шукана схема має складатися з елементів, які відповідають елементарним булевим функціям функціонально повної системи E:

(а) $\boldsymbol{a}_f = (0110)$, $E = \{\rightarrow, \neg\}$;

(б) $\boldsymbol{a}_f = (10111000)$, $E = \{\wedge, \neg\}$;

(в) $\boldsymbol{a}_f = (0101110001111001)$, $E = \{\oplus, \wedge, 1\}$;

(г) $\boldsymbol{a}_f = (1001111111110110)$, $E = \{\oplus, \vee, 1\}$.

2. Побудувати якомога простішу комбінаційну схему над E_0, яка реалізує булеву функцію f, задану вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (1101)$;

(б) $\boldsymbol{a}_f = (11011001)$;

(в) $\boldsymbol{a}_f = (0001111100110110)$.

3. Для функції, заданої формулою A, побудувати комбінаційну схему над E_0, що реалізує цю функцію та має складність не більшу п'яти:

(а) $A = (a \vee bc)\big((a \vee b)(\neg b \vee d) \vee (c \vee \neg d)(\neg a \neg bc \vee de)\big)$;

(б) $A = (a \vee b)\big((ab \vee \neg b \neg c)f \vee (\neg b \vee \neg c \neg e)(\neg b \vee \neg c) \vee (\neg a \neg b \vee ab \vee f)c \vee \vee ab \vee ef\big)$.

4. Довести, що для довільної булевої функції f складність оптимальної РК схеми, яка реалізує f, не перевищує складності будь-якої комбінаційної схеми над E_0, що реалізує f.

5. Довести, що коли для комбінаційної схеми зі складністю m над E_0, яка реалізує булеву функцію $f(x_1, x_2, \ldots, x_n)$, виконується $m > n\,2^{n-1}$, то ця схема не оптимальна.

6.15. Проблема мінімізації формул алгебри логіки

В обох розглянутих у попередніх розділах алгебрах схем умови роботи схем задають за допомогою булевих функцій. Основа для розв'язання однієї з центральних задач у цих алгебрах — задачі синтезу — це формули, які зображують відповідні булеві функції. Задачу синтезу й іншу важливу задачу алгебр схем — задачу оптимізації — не розв'язують окремо одна від одної. Зазвичай розглядають задачу оптимального синтезу, що полягає у проектуванні схеми, яка задовольняє задані умови роботи і водночас є оптимальною за певними критеріями. На практиці використовують такі критерії: складність, вартість (ціна), надійність, швидкодія, енергоспоживання схеми тощо. Найпоширеніші критерії — це вартість схеми і безпосередньо пов'язаний із вартістю показник складності схеми.

Вартістю (***ціною***) схеми S називається величина $C = \sum_{i=1}^{r} c_i n_i$, де r — кількість типів елементів у системі E (базових елементів), c_i — ціна елемента i-го типу і n_i — кількість елементів i-го типу, використаних у схемі S, $i = 1, 2, \ldots, r$.

Часто вважають, що ціни всіх базових елементів приблизно однакові, тобто різницею в цінах цих елементів можна знехтувати. У такому разі вартість схеми визначається загальною кількістю її елементів, яку називають ***складністю*** схеми. Надалі як критерій оптимальності синтезованої схеми буде використано саме складність схеми.

Оскільки складність схеми безпосередньо залежить від параметрів формули, за якою синтезують цю схему, розглянемо відповідні характеристики для формул алгебри логіки.

Мінімальною формою для формули алгебри логіки, або ***мінімальною формулою***, називають формулу, рівносильну даній формулі і таку, що має найменшу можливу складність. Складність формули оцінюють за кількома параметрами. Наприклад, складністю формули можна вважати кількість усіх її операндів. Іншим критерієм складності формули може бути загальна кількість операцій у ній. Першим із цих критеріїв зручно користуватися для оцінки складності релейно-контактної схеми, синтезованої за даною формулою. Другий показник характеризує складність для відповідної комбінаційної схеми.

Часто, визначаючи складність комбінаційних схем, нехтують елементами заперечення і рахують тільки комбінаційні елементи, які реалізують бінарні булеві операції. Неважко переконатися, що тоді кількість операндів на одиницю перевищує кількість бінарних операцій у формулі. Отже, обидва розглянуті критерії складиості формул стають практично еквівалентними. Це дає змогу обрати основним критерієм складності формули кількість її операндів.

Уведений показник не треба плутати з кількістю аргументів у формулі. Наприклад, формула $xyz \vee \overline{x}y\overline{z} \vee x\overline{y} \vee \overline{x}$ містить дев'ять операндів і залежить від трьох аргументів x, y і z. Отже, складність цієї формули дорівнює 9.

Підсумовуючи вищезазначене, задачу оптимального синтезу схеми можна сформулювати так. Задано умови роботи шуканої схеми у вигляді булевої функції f. Крім того, задано набір базових булевих функцій Σ: у випадку релейно-контактних схем це функції диз'юнкції, кон'юнкції та заперечення, а для комбінаційних схем — система функцій Φ_k. Потрібно визначити мінімальну формулу над Σ, яка зображує задану булеву функцію f. З огляду на важливість сформульовану задачу виділяють в окрему проблему, яка дістала назву ***мінімізації формул алгебри логіки***.

Задачу мінімізації формули можна в принципі розв'язати ***методом перебору***. Потрібно послідовно перебирати всі формули над Σ, які містять один операнд, відтак всі формули над Σ, що містять два операнди, три операнди і т. д., а потім для кожної з формул слід перевірити, чи реалізує вона задану булеву функцію f. Перша з формул, для якої відповідь позитивна, є мінімальною.

Процес перебору обов'язково завершиться через певну скінченну кількість кроків. Гарантією цього є, по-перше, існування формули F над Σ, що реалізує булеву функцію f, а по-друге — той очевидний факт, що множина всіх різних формул над Σ, складність яких не перевищує складності формули F і змінні яких вибирають з певної скінченної множини символів, є скінченною множиною.

Однак запропонований алгоритм настільки складний і громіздкий, що практичне його застосування нереальне.

Задачу побудови більш ефективних і зручних методів мінімізації формул алгебри логіки в загальному вигляді, тобто для довільної системи базових булевих функцій Σ, досі не розв'язано. Є всі підстави вважати, що таких універсальних методів мінімізації формул, які б істотно відрізнялися від описаного методу перебору, не існує взагалі.

Тому на практиці основну увагу приділяють розробці методів мінімізації для окремих класів формул. Зокрема, на сьогодні найбільш досліджено та вивчено проблему побудови мінімальних формул у класах диз'юнктивних нормальних і кон'юнктивних нор-мальних форм. Обмежимося викладенням основних (класичних) методів мінімізації диз'юнктивних нормальних форм, оскільки методи мінімізації кон'юнктивних нормальних форм подібні.

Історично задача мінімізації формул у класі диз'юнктивних нормальних форм дістала назву ***мінімізації булевих функцій***. Така назва фігурує в багатьох підручниках, монографіях і статтях, присвячених цій проблематиці. Однак необхідно підкреслити некоректність цього терміна. Насправді мова йде про мінімізацію (тобто спрощення) лише формули, яка зображує певну булеву функцію, а сама ця функція залишається незмінною протягом усієї процедури розв'язання задачі.

6.16. Мінімальні, скорочені та тупикові ДНФ булевих функцій

Нагадаємо, що елементарною кон'юнкцією називається кон'юнкція деяких змінних або їх заперечень, в якій кожна змінна зустрічається не більше одного разу; диз'юнктивна нормальна форма (ДНФ) — це формула, що має вигляд диз'юнкції елементарних кон'юнкцій, а ДНФ булевої функції f називається ДНФ, яка її зображує (див. розд. 6.5).

Для будь-якої булевої функції f ($f \neq 0$) існує її ДНФ (наприклад, ДДНФ). Однак на відміну від єдиної ДДНФ функції f може існувати, взагалі кажучи, кілька різних її ДНФ.

Приклад 6.5. Кожна з ДНФ $\overline{x}\overline{y} \vee x\overline{y} \vee xy$, $\overline{y} \vee xy$, $\overline{x}\overline{y} \vee x$, $x \vee \overline{y}$ зображує ту саму булеву функцію $f(x, y)$.

Мінімальною диз'юнктивною нормальною формою (***МДНФ***) функції f називається ДНФ, яка містить найменшу можливу кількість операндів серед усіх ДНФ функції f.

Задача побудови МДНФ для булевої функції f називається ***канонічною задачею мінімізації***.

Іноді розглядають задачу побудови ДНФ функції f, яка містить найменшу можливу кількість елементарних кон'юнкцій; цю ДНФ називають ***найкоротшою*** (***НДНФ***). Така задача виникає тоді, коли для синтезу комбінаційних схем використовують комбінаційні елементи, що реалізують одночасно кон'юнкцію довільної кількості булевих змінних (двох і більше), а складність схеми вимірюють, як і раніше, кількістю її елементів.

Усі методи мінімізації в класі ДНФ грунтуються на понятті імпліканти булевої функції.

Нагадаємо, що через N_f позначено множину всіх двійкових кортежів довжини n, на яких n-арна булева функція f набуває значення 1, тобто $N_f = \{\sigma \mid f(\sigma) = 1, \sigma\in B^n\}$. Елементи множини N_f будемо називати ***одиницями*** функції f.

Кажуть, що булева функція g ***імплікує*** булеву функцію f, якщо $N_g \subseteq N_f$.

Імплікантою булевої функції f називається елементарна кон'юнкція, яка її імплікує.

Нехай K — деяка імпліканта булевої функції f. Оскільки $N_K \subseteq N_f$, то говоритимемо, що імпліканта K *накриває* деякі *одиниці* функції f.

Із властивостей диз'юнкції випливає, що диз'юнкція будь-якого числа імплікант функції f імплікує функцію f. Якщо ж імпліканти K_1, K_2, …, K_r у сукупності накривають усі одиниці функції f, тобто $N_f = \bigcup_{i=1}^{r} N_{K_i}$ (зрозуміло, що не може бути $N_f \subset \bigcup_{i=1}^{r} N_{K_i}$), то диз'юнкція $\bigvee_{i=1}^{r} K_i$ є ДНФ функції f.

Зрозуміле й обернене твердження: будь-яка елементарна кон'юнкція з ДНФ функції f є її імплікантою і накриває деякі одиниці функції f, а всі члени (елементарні кон'юнкції) ДНФ функції f у сукупності накривають усі одиниці функції f.

Елементарна кон'юнкція, що містить n операндів, тобто елементарна кон'юнкція максимальної довжини n, може накривати лише одну єдину одиницю n-арної булевої функції f. Водночас елементарна кон'юнкція довжини k $(1 \le k \le n)$ може накривати 2^{n-k} одиниць функції f. Це пояснюється тим, що для елементарної кон'юнкції K довжини k $n - k$ її змінних є неістотними, тому елементарна кон'юнкція K набуває одного й того самого значення на 2^{n-k} двійкових кортежах довжини n.

Таким чином, зменшення довжини елементарної кон'юнкції K, яка є імплікантою і накриває деякі одиниці булевої функції f, тобто вилучення з K одного чи більшого числа операндів зумовлює те, що одержана елементарна кон'юнкція K' накриває більшу кількість двійкових кортежів. Якщо ж при цьому K' — імпліканта функції f, то вона обов'язково накриває всі ті одиниці функції f, які накриває елементарна кон'юнкція K, а також ще деякі інші одиниці функції f. Інакше кажучи, $N_K \subset N_{K'} \subseteq N_f$. Звідси, зокрема, випливає, що всі імпліканти булевої функції f є частинами елементарних кон'юнкцій, які утворюють ДДНФ функції f. Ця властивість лежить в основі всіх методів мінімізації ДНФ.

Імпліканта булевої функції f називається ***простою імплікантою***, якщо після вилучення з неї будь-якого операнда одержана елементарна кон'юнкція вже не є імплікантою функції f.

Отже, неможливо зменшити довжину простої імпліканти без того, щоб вона не перестала бути імплікантою функції f. Наголосимо, що всі прості імпліканти функції f є частинами елементарних кон'юнкцій, які складають ДДНФ функції f.

Нехай F — формула, що зображує булеву функцію f, а G — формула, що зображує булеву функцію g, яка імплікує функцію f (наприклад, G — імпліканта функції f). Тоді формула $G \to F$ (читається "G імплікує F") тотожно дорівнює 1. Справді, формула $G \to F$ може набувати значення 0 на деякому наборі значень аргументів лише тоді, коли на цьому наборі формула G набуває значення 1, а формула F — 0. Але це неможливо за означенням поняття "g імплікує f". Сказане пояснює вибір термінів "імплікує" й "імпліканта".

Приклад 6.6. Розглянемо булеву функцію $f(x, y, z)$, ДДНФ якої — $\bar{x}\bar{y}z \vee \bar{x}y\bar{z} \vee \bar{x}yz \vee x\bar{y}\bar{z} \vee x\bar{y}z \vee xy\bar{z}$ Тоді $N_f = \{(0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0)\}$. Для булевої функції $g(x, y, z)$, яку зображує формула $x\bar{y} \vee x\bar{z}$, $N_g = \{(1, 0, 0), (1, 0, 1), (1, 1, 0)\}$. Отже, функція g імплікує функцію f. Імплікантами функції f є, наприклад, елементарні кон'юнкції $\bar{x}\bar{y}z$, $xy\bar{z}$, $x\bar{y}$, $x\bar{z}$, $y\bar{z}$. Прості імпліканти функції f —

$\bar{x}z$, $\bar{y}z$, $\bar{x}y$, $y\bar{z}$, $x\bar{y}$ та $x\bar{z}$. Неважко переконатися (за допомогою простого перебору), що це всі прості імпліканти функції f.

Важливість поняття простої імпліканти булевої функції можна обгрунтувати за допомогою двох таких теорем.

Теорема 6.9. Диз'юнкція всіх простих імплікант булевої функції f зображує цю функцію.

Доведення. Нехай K_1, K_2, ..., K_m — усі прості імпліканти n-арної булевої функції f. Необхідно довести, що диз'юнкція $D = \bigvee\limits_{i=1}^{r} K_i$ набуває значення 1 на тих і тільки тих двійкових кортежах довжини n, на яких функція f набуває значення 1, тобто що $N_D = N_f$. Включення $N_D \subseteq N_f$ випливає безпосередньо з означення поняття імпліканти та властивостей операції диз'юнкції.

Навпаки, нехай для якогось двійкового кортежу $\sigma \in B^n$ виконується $f(\sigma) = 1$, тобто $\sigma \in N_f$. Тоді в ДДНФ функції f є елементарна кон'юнкція K, яка накриває цю одиницю функції f, тобто $K(\sigma) = 1$. Оскільки для елементарної кон'юнкції K існує принаймні одна проста імпліканта K_i, яка є частиною K (можливо, що $K = K_i$), то імпліканта K_i, а отже, і вся диз'юнкція D набувають значення 1 на наборі σ, тобто $\sigma \in N_{K_i} \subseteq N_D$. Тому $N_f \subseteq N_D$.

Диз'юнкція всіх простих імплікант булевої функції f.називається ***скороченою диз'юнктивною нормальною формою (СДНФ)*** функції f.

СДНФ, узагалі кажучи, простіша за ДДНФ. Проте інколи СДНФ збігається з ДДНФ. Наприклад, для булевої функції $f(x, y)$, ДДНФ якої — $x\bar{y} \vee \bar{x}y$, простими імплікантами є $x\bar{y}$ і $\bar{x}y$, тому СДНФ цієї функції збігається з її ДДНФ.

Однак у більшості випадків скорочену ДНФ D булевої функції f можна спростити внаслідок того, що деякі з простих імплікант функції f накривають ті одиниці функції f, які накриваються в сукупності кількома іншими імплікантами цієї функції. Усі такі *зайві* прості імпліканти можна вилучити з формули D. Одержану ДНФ називають *тупиковою*.

Отже, диз'юнкція простих імплікант, яка зображує булеву функцію f і в якій немає зайвих імплікант, називається ***тупиковою диз'юнктивною нормальною формою*** функції f.

Приклад 6.7. Для булевої функції $f(x, y, z)$ із прикладу 6.6 СДНФ має вигляд $D = \bar{x}z \vee \bar{y}z \vee \bar{x}y \vee y\bar{z} \vee x\bar{y} \vee x\bar{z}$.

Проста імпліканта $x\bar{y}$ накриває кортежі (1, 0, 0) і (1, 0, 1). Але ці самі кортежі накривають відповідно імпліканти $x\bar{z}$ та $\bar{y}z$. Отже, $x\bar{y}$ — зайва імпліканта, вилучивши яку з формули D, дістанемо рівносильну формулу $D_1 = \bar{x}z \vee \bar{y}z \vee \bar{x}y \vee y\bar{z} \vee x\bar{z}$. Аналогічно з формули D_1 можна вилучити імпліканту $\bar{x}y$, що накриває кортежі (0, 1, 0) і (0, 1, 1), які, у свою чергу, накриваються відповідно імплікантами $y\bar{z}$ та $\bar{x}z$. Одержимо рівносильну формулу $T_1 = \bar{x}z \vee \bar{y}x \vee y\bar{z} \vee x\bar{z}$. Легко перевірити, що в останній формулі вже немає зайвих імплікант. Отже, T_1 — тупикова ДНФ функції f.

Порядок вилучення зайвих імплікант зі скороченої ДНФ функції f може бути й іншим. При цьому отримаємо, взагалі кажучи, іншу тупикову ДНФ. Наприклад, вилучимо імпліканту $\bar{x}z$, що накриває кортежі (0, 1, 1) і (0, 0, 1), які накриваються відповідно імплікантами $\bar{y}z$ і $\bar{x}y$. Одержимо $D_1' = \bar{y}z \vee \bar{x}y \vee y\bar{z} \vee x\bar{y} \vee x\bar{z}$.

Далі, імпліканта $y\bar{z}$ накриває кортежі (0, 1, 0) і (1, 1, 0), що накриваються також відповідно імплікантами $\bar{x}y$ і $x\bar{z}$. Маємо $D_2' = \bar{y}z \vee \bar{x}y \vee x\bar{z} \vee x\bar{y}$.

Відтак в останній формулі можна вилучити імпліканту $x\bar{z}$, бо вона накриває кортежі (1, 0, 0) і (1, 0, 1), які накриваються також відповідно імплікантами $x\bar{z}$та $\bar{y}z$. Отже, остаточно маємо $T_2 = \bar{y}z \vee \bar{x}y \vee x\bar{z}$.

У формулі T_2 вже немає зайвих імплікант, і тому вона є тупиковою ДНФ функції f. Існують й інші її тупикові ДНФ. Неважко переконатися, що, наприклад, формула $T_3 = \bar{x}z \vee x\bar{y} \vee y\bar{z}$ є також тупиковою ДНФ функції f.

Зауважимо, що одержані тупикові ДНФ мають різні складності: складності формул T_2 і T_3 збігаються й дорівнюють 6, а складність формули T_1 дорівнює 8.

Теорема 6.10. Мінімальна ДНФ булевої функції f, яка не дорівнює тотожно 0 чи 1, є диз'юнкцією деяких простих імплікант функції f.

Доведення. Розглянемо довільну МДНФ n-арної функції f:

$$M = K_1 \vee K_2 \vee \ldots \vee K_m. \tag{6.11}$$

Припустімо, що якась елементарна кон'юнкція з M, скажімо K_1, не є простою імплікантою. Тоді з неї шляхом вилучення деяких операндів можна утворити просту імпліканту K_1'.

Доведемо, що формула $M' = K_1' \vee K_2 \vee \ldots \vee K_m$ зображує функцію f. Оскільки K_1' є частиною K_1, то K_1' накриває всі ті одиниці функції f, які накриває K_1. Отже, для довільного двійкового кортежу $\sigma \in B^n$ із того, що значення формули M на кортежі σ дорівнює 1 (тобто $f(\sigma) = 1$), випливає, що й значення M' на σ дорівнює 1.

Навпаки, нехай для кортежу $\sigma \in B^n$ формула M' дорівнює 1. Це означає, що або K_1', або $K_2 \vee K_3 \vee \ldots \vee K_m$ на кортежі σ дорівнюють 1 (обидві ці умови можуть виконуватися одночасно). Тоді співвідношення $f(\sigma) = 1$ для першого випадку випливає з того, що K_1' — імпліканта функції f, а для другого — з того, що формула M зображує функцію f і має вигляд (6.11).

Отже, формула M' рівносильна формулі M і має меншу складність. Це суперечить припущенню про мінімальність формули M. Теорему доведено.

Наслідок 6.10.1. Будь-яка мінімальна ДНФ функції f є її тупиковою ДНФ.

Справді, нехай M — мінімальна ДНФ функції f. Тоді за теоремою 6.10 вона є диз'юнкцією деяких простих імплікант функції f. Якщо припустити, що M не є тупиковою ДНФ, то, вилучивши з неї зайві прості імпліканти, дістанемо ДНФ M' функції f, складність якої менша, ніж складність формули M. Це суперечить тому, що M — мінімальна ДНФ функції f.

Отже, усі мінімальні ДНФ функції f містяться серед її тупикових ДНФ. Крім того, зауважимо, що, як показано в прикладі 6.7, не кожна тупикова ДНФ функції f є мінімальною.

6.15. Задачі і вправи

1. Знайти серед елементарних кон'юнкцій множини K імпліканти булевої функції f, заданої вектором значень $\boldsymbol{a}_f$:

(а) $K = \{x, \bar{z}, xy, y\bar{z}\}$, $\boldsymbol{a}_f = (00101111)$;

(б) $K = \{x\bar{y}, y\bar{z}, x, xy\bar{z}\}$, $\boldsymbol{a}_f = (01111110)$;

(в) $K = \{x\bar{z}, y\bar{z}, x\overline{zw}, x\bar{w}, xy\bar{z}w\}$, $\boldsymbol{a}_f = (1011111011001111)$.

2. Із даної множини елементарних кон'юнкцій K виділити прості імпліканти функції f, заданої вектором значень $\boldsymbol{a}_f$:

(а) $K = \{x\bar{y}, y\bar{z}, x, \bar{x}yz\}$, $\boldsymbol{a}_f = (01111110)$;

(б) $K = \{\bar{z}w, y\bar{z}, x\bar{z}, xz, xy, \overline{xy}z\}$, $\boldsymbol{a}_f = (1011110011001111)$.

3. Довести, що жодна проста імпліканта монотонної функції не містить заперечень змінних.

4. Визначити число булевих функцій $f(x_1, x_2, \ldots, x_n)$, для яких задана елементарна кон'юнкція K рангу k є імплікантою.

5. Довести, що довільна булева функція від n змінних може бути реалізована ДНФ, яка містить не більше ніж 2^{n-1} елементарних кон'юнкцій.

6. Довести, що число елементарних кон'юнкцій у скороченій ДНФ довільної булевої функції $f(x_1, x_2, \ldots, x_n)$ не перевищує $|N_f|\big(|N_f| + 1\big)/2$.

7. Нехай K, K_1, K_2 — різні імпліканти булевої функції $f(x_1, x_2, \ldots, x_n)$. Довести, що $K = K_1 \vee K_2$ тоді й тільки тоді, коли існують така змінна x_i та значення $a \in B$, що $K_1 = x_i^a K$ і $K_2 = x_i^{\neg a} K$.

8. Нехай K, K_1, K_2 — різні елементарні кон'юнкції зі скороченої ДНФ, а k, k_1, k_2 — відповідні ранги цих кон'юнкцій. Довести, що коли $K \vee K_1 \vee K_2 = K_1 \vee K_2$, то $k_1 + k_2 \geq k + 2$.

6.17. Метод Квайна — Мак-Класкі побудови скороченої ДНФ булевої функції

Теорема 6.10 та її наслідок лежать в основі загальної схеми розв'язання канонічної задачі мінімізації формул алгебри логіки. Ця схема складається з двох основних етапів.

Перший етап. Знайти всі прості імпліканти заданої булевої функції f, тобто побудувати її скорочену ДНФ.

Другий етап. Вилучивши зайві прості імпліканти, побудувати всі тупикові ДНФ функції f, серед яких вибрати мінімальні ДНФ.

Є різні методи побудови скороченої ДНФ для заданої булевої функції. Розглянемо метод, який на початку 50-х років XX століття запропонував американський математик У. Квайн, а згодом удосконалив інший його співвітчизник Е. Мак-Класкі.

Щоб зрозуміти сутність цього методу, пригадаймо процедуру зведення булевої формули F, яка зображує функцію f, до ДДНФ цієї функції (див. доведення теореми 6.4). Після пункту 5 цієї процедури дістаємо деяку ДНФ D функції f. У пункті 6 ДНФ D перетворюється у ДДНФ за допомогою так званого *правила розщеплення* (*розгортання*)

$$A = Ax \vee A\bar{x}, \tag{6.12}$$

де A — елементарна кон'юнкція з D (імпліканта функції f), а x — змінна функції f, яка не входить в A.

Розглянемо детальніше цей процес. Нехай $H = \bigvee_{i=1}^{m} K_i$ — шукана СДНФ, а K_i — якась проста імпліканта n-арної булевої функції f. Для зручності будемо вважати, що K_i дорівнює кон'юнкції перших p змінних функції f, тобто $K_i = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_p^{\alpha_p}$. На всіх 2^{n-p} двійкових кортежах довжини n із множини $N_{K_i} = \{(\alpha_1, \alpha_2, \ldots, \alpha_p, \sigma_{p+1}, \ldots, \sigma_n) \mid \sigma_j \in B, j = p+1, p+2, \ldots, n\}$ елементарна кон'юнкція K_i набуває значення 1. Із того, що K_i — імпліканта функції f, випливає, що й функція f набуває значення 1 на всіх кортежах з N_{K_i}. Отже, у ДДНФ функції f входять 2^{n-p} елементарних кон'юнкцій виду $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_p^{\alpha_p} x_{p+1}^{\sigma_{p+1}} \ldots x_n^{\sigma_n}$, $\sigma_j \in B$, $j = p+1, p+2, \ldots, n$.

Усі ці елементарні кон'юнкції (точніше, підформулу ДДНФ, яка містить усі ці елементарні кон'юнкції) дістанемо в результаті послідовного застосування до K_i формули розгортання (6.12). На першому кроці замінимо у формулі H елементарну кон'юнкцію K_i на пару $K_i x_{p+1}$ і $K_i \overline{x_{p+1}}$, потім кожну з одержаних елементарних кон'юнкцій — на пари $K_i x_{p+1} x_{p+2}$, $K_i x_{p+1} \overline{x_{p+2}}$ та $K_i \overline{x_{p+1}} x_{p+2}$, $K_i \overline{x_{p+1}}\, \overline{x_{p+2}}$ і т. д. Застосувавши аналогічну процедуру до всіх простих імплікант з H, одержимо рівносильну H формулу G, яка містить усі елементарні кон'юнкції, що складають ДДНФ функції f.

Відзначимо, що оскільки різні прості імпліканти з H можуть накривати одні й ті самі одиниці функції f, то у формулі G деякі з елементарних кон'юнкцій ДДНФ можуть зустрічатися більше одного разу. Користуючись властивістю ідемпотентності диз'юнкції, вилучаємо з формули G всі елементарні кон'юнкції, які повторюються. Одержуємо ДДНФ функції f.

Зрозуміло, що, як і всі рівносильні перетворення, описаний процес має властивість оборотності, тобто, маючи ДДНФ функції f і послідовно застосовуючи до неї співвідношення (6.12) у вигляді так званої ***формули склеювання*** (за змінною x)

$$Ax \vee A\overline{x} = A, \tag{6.13}$$

можна побудувати скорочену ДНФ функції f.

Певні труднощі виникають через те, що для реалізації склеювання нам потрібна, взагалі кажучи, не ДДНФ функції f, а рівносильна їй формула G, що відрізняється від ДДНФ тим, що деякі члени ДДНФ у ній можуть повторюватися кілька разів. Щоб усунути цю проблему, замість формули склеювання (6.13) слід застосовувати ***формулу неповного склеювання***

$$Ax \vee A\overline{x} = A \vee Ax \vee A\overline{x}. \tag{6.14}$$

У результаті застосування неповного склеювання всі пари членів формули, для яких було виконано склеювання, продовжують залишатись у формулі, а тому їх можна використовувати в подальших склеюваннях з іншими членами цієї формули. Щоб уникнути повторного склеювання одних і тих самих пар членів, ці пари доцільно якось позначати.

Проаналізуємо основну формулу всіх виконуваних рівносильних перетворень — формулу (6.14). Нехай G — якась ДНФ функції f (зокрема, її ДДНФ), а Ax, $A\overline{x}$ — її члени. Тоді Ax і $A\overline{x}$ — імпліканти функції f. Із формули (6.13) випливає, що A — також імпліканта функції f. Якщо Ax і $A\overline{x}$ накривають відповідно множини двійкових кортежів L_1 і L_2, то A накриває множину $L_1 \cup L_2$. Тому, виконуючи рівносильні перетворення за формулою неповного склеювання (6.14), ми кожного разу додаємо до ДНФ функції f нову, "простішу", імпліканту. Отже, якщо взяти ДДНФ функції f і виконати всі можливі склеювання, одержимо формулу D', яка є диз'юнкцією *всіх* імплікант функції f. Серед інших формула D' містить і всі прості імпліканти даної функції.

Щоб отримати з формули D' СДНФ функції f, послідовно застосуємо до елементарних кон'юнкцій формули D' ***формулу елементарного поглинання***

$$Ax^{\alpha} \vee A = A. \tag{6.15}$$

Смисл цієї формули полягає в тому, що з двох імплікант Ax^{α} й A, одна з яких є частиною іншої, залишаємо у формулі тільки "простішу" імпліканту. Кажемо, що імпліканта A *поглинає* імпліканту Ax^{α}.

Після виконання всіх поглинань одержимо шукану СДНФ функції f.

В описаній процедурі побудови СДНФ спочатку виконувались усі неповні склеювання, а тоді всі елементарні поглинання. Зручніше виконувати ці перетворення почергово у певному порядку, який і встановлює ***метод Квайна***.

Для заданої n-арної булевої функції f спочатку будуємо її ДДНФ F_0, а потім послідовність F_0, F_1, F_2, … ДНФ функції f. Перетворити формулу F_i цієї послідовності у формулу F_{i+1} можна за два кроки:

а) у формулі F_i виконати всі неповні склеювання для елементарних кон'юнкцій довжини $n - i$,

б) вилучити з одержаної формули всі ті елементарні кон'юнкції довжини $n-i$, які поглинаються "простішими" елементарними кон'юнкціями за формулою (6.15), $i = 0, 1, 2, \ldots$.

Процес побудови послідовності F_0, F_1, F_2, ... закінчується, коли в деякій формулі F_k не можна виконати жодного склеювання. Формула F_k є шуканою скороченою ДНФ функції f.

Зрозуміло, що метод Квайна приводить до бажаного результату не більше ніж за n кроків.

Приклад 6.8. Продемонструємо роботу методу Квайна для булевої функції $f(x, y, z)$, що задана своєю ДДНФ

$$F_0 = \bar{x}\bar{y}\bar{z} \vee \bar{x}\bar{y}z \vee \bar{x}yz \vee x\bar{y}\bar{z} \vee x\bar{y}z \vee xy\bar{z}.$$

Описуючи та пояснюючи процедуру реалізації методу Квайна, для ідентифікації імплікант (елементарних кон'юнкцій) функції f скористаємося частково системою позначень, запропонованою Мак-Класкі. Елементарну кон'юнкцію позначатимемо впорядкованою множиною номерів тих двійкових кортежів, які накриваються нею.

На першому кроці виконуємо неповні склеювання для елементарних кон'юнкцій формули F_0. Спочатку склеїмо імпліканти {0} і {1}, {4} і {5} за змінною z; одержимо нові імпліканти: $\bar{x}\bar{y}(\{0, 1\})$, $x\bar{y}(\{4, 5\})$. Далі, склеїмо імпліканти {1} і {3}, {4} і {6} за змінною y, а імпліканти {0} і {4}, {1} і {5} — за змінною x. Одержимо відповідно нові імпліканти: $\bar{x}z$ $(\{1, 3\})$, $x\bar{z}$ $(\{4, 6\})$ та $\bar{y}\bar{z}$ $(\{0, 4\})$, $\bar{y}z$ $(\{1, 5\})$.

Оскільки всі без винятку елементарні кон'юнкції довжини 3 брали участь у склеюваннях, то всі вони поглинаються відповідними "простішими" елементарними кон'юнкціями довжини 2.

Імпліканта, позначена множиною номерів P, поглинається імплікантою, що позначена множиною номерів R, тоді й тільки тоді, коли $P \subseteq R$.

Таким чином, після першого кроку одержимо формулу

$$\begin{array}{cccccc} \{0, 1\} & \{4, 5\} & \{1, 3\} & \{4, 6\} & \{0, 4\} & \{1, 5\} \end{array}$$
$$F_1 = \bar{x}\bar{y} \vee x\bar{y} \vee \bar{x}z \vee x\bar{z} \vee \bar{y}\bar{z} \vee \bar{y}z.$$

У формулі F_1 склеїмо імпліканти {0, 4} та {1, 5} за змінною z, а імпліканти {0, 1} і {4, 5} — за змінною x. В обох цих випадках з'явиться нова імпліканта $\bar{y}$ позначена множиною {0, 1, 4, 5}, яка поглинає всі чотири імпліканти довжини 2, що брали участь в її утворенні. Імпліканти {1, 3} і {4, 6} не було використано у склеюваннях, і тому вони залишаються. Отже, маємо

$$\begin{array}{ccc} \{1, 3\} & \{4, 6\} & \{0, 1, 4, 5\} \end{array}$$
$$F_2 = \bar{x}z \vee x \vee \bar{y}.$$

В одержаній формулі неможливі жодні склеювання, тому вона є шуканою СДНФ. Легко перевірити, що серед імплікант формули F_2 немає зайвих, із чого випливає, що F_2 — мінімальна ДНФ функції $f(x, y, z)$.

Неважко переконатися, що при склеюванні двох імплікант однакової довжини, позначених множинами номерів P і R, одержимо нову "простішу" імпліканту, якій відповідає множина номерів $P \cup R$.

Застосування методу Квайна для мінімізації ДНФ булевої функції від великої кількості змінних потребує чимало зусиль. Труднощі виникають у зв'язку зі збільшенням кількості членів у ДНФ, а отже, зі значним збільшенням варіантів склеювання імплікант.

Для зменшення цих труднощів Мак-Класкі запропонував модифікований і вдосконалений алгоритм.

На відміну від методу Квайна, алгоритм Мак-Класкі оперує не з елементарними кон'юнкціями, а з вищезгаданими множинами номерів відповідних двійкових кортежів та з деякими характеристиками цих множин.

Як уже було зазначено, Мак-Класкі запропонував позначати будь-яку імпліканту K довжини m n-арної булевої функції $f(x_1, x_2, \ldots, x_n)$ упорядкованою множиною з 2^{n-m} номерів тих двійкових кортежів із B^n, які накриває імпліканта K. Крім того, кожній імпліканті K поставимо у відповідність так званий індекс або вектор індексів імпліканти K.

Індексом імпліканти K довжини n назвемо кількість одиниць (тобто вагу, або норму) двійкового кортежу, який накриває імпліканта K. Наприклад, для $n = 3$ індекс імпліканти $x\bar{y}z$ дорівнює 2, а індекси імплікант $\bar{x}\bar{y}\bar{z}$ і $\bar{x}\bar{y}z$ — відповідно 0 і 1.

Якщо імпліканта K має довжину m $(1 \le m < n)$ і змінні $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ $(k = n - m)$ не входять (із запереченнями або без заперечень) в імпліканту K, то їй ставлять у відповідність ***вектор індексів*** $(2^{n-i_1}, 2^{n-i_2}, \ldots, 2^{n-i_k})$.

Сформулюємо без доведення алгоритм Мак-Класкі для реалізації неповних склеювань і елементарних поглинань у процесі побудови СДНФ заданої n-арної булевої функції f. Обгрунтування цього алгоритму можна знайти в монографії В. М. Глушкова [8].

Алгоритм Мак-Класкі

1. Для того щоб імпліканти K_1 і K_2 довжини n склеювалися між собою, необхідно й достатньо, щоб було виконано такі умови:

а) їх індекси відрізняються на одиницю;

б) їх номери відрізняються на степінь числа 2;

в) індекс імпліканти з більшим номером має бути більшим від індекса імпліканти з меншим номером.

Результат склеювання позначатимемо парою номерів (у фігурних дужках) імплікант K_1 і K_2. Поруч будемо записувати відповідний вектор індексів, який міститиме одну компоненту, що дорівнює модулю різниці номерів імплікант K_1 і K_2.

2. Імпліканта K_1, позначена множиною номерів P_1, є частиною імпліканти K_2, позначеної множиною номерів P_2, тоді й тільки тоді, коли $P_1 \subseteq P_2$. Отже, імпліканта K_1 поглинається імплікантою K_2 тоді й лише тоді, коли $P_1 \subseteq P_2$.

3. Дві імпліканти K_1 і K_2 довжини m ($1 \le m < n$) склеюються між собою тоді й тільки тоді, коли:

а) у них однакові вектори індексів;

б) найменші (перші) номери l_1 і l_2 у множинах номерів P_1 і P_2, якими позначено імпліканти K_1 і K_2, є номерами тих імплікант довжини n, які склеюються між собою.

Результату склеювання відповідає множина номерів $P_1 \cup P_2$. Для зручності будемо впорядковувати номери в новоутворених множинах. Вектор індексів, що відповідає новій імпліканті, одержимо з вектора індексів імпліканти K_1 або K_2, додаючи до нього компоненту $|l_1 - l_2|$ так, щоб компоненти нового вектора було впорядковано за зростанням.

Основою алгоритму Мак-Класкі є метод Квайна, тобто процедура побудови послідовності ДНФ $F_0, F_1, F_2, \ldots$ функції f за допомогою застосування операцій неповного склеювання й елементарного поглинання. Змінено й удосконалено лише способи виконання та запису результатів цих операцій.

Приклад 6.9. Нехай булева функція $f(x, y, z, w)$ набуває значення 1 на двійкових кортежах із номерами 0, 1, 3, 4, 6, 8, 9, 11, 12, 13. Запишемо процес побудови СДНФ функції f за алгоритмом Мак-Класкі (див.табл. 6.7).

Для зручності згрупуємо всі імпліканти довжини 4 в першому стовпчику таблиці за їх індексами. Індекс будемо записувати праворуч від фігурних дужок з номером імпліканти.

Таблиця 6.7

0	1	2
{0} 0 *	{0, 1}(1) *	
	{0, 4}(4) *	
{1} 1 *	{0, 8}(8) *	
{4} 1 *	{1, 3}(2) *	{0, 1, 8, 9} (1, 8)
{8} 1 *	{1, 9}(8) *	
{3} 2 *	{4, 6}(2)	{0, 4, 8, 12} (4, 8)
{6} 2 *	{4, 12}(8) *	
{9} 2 *	{8, 9}(1) *	{1, 3, 9, 11} (2, 8)
{12} 2 *	{8, 12}(4) *	
{11} 3 *	{3, 11}(8) *	{8, 9, 12, 13} (1, 4)
{13} 3 *	{9, 11}(2) *	
	{9, 13}(4) *	
	{12, 13}(1) *	

Аналогічно в наступних стовпчиках також згрупуємо імпліканти згідно з їх індексами, які відповідають першим номерам у множині номерів імпліканти. Імпліканти, що брали участь принаймні в одному склеюванні, позначатимемо в таблиці зірочками. Вектор індексів імпліканти записуватимемо праворуч від множини її номерів. Зауважимо, що склеювання різних імплікант може давати один і той самий результат. Вписуємо його в таблицю тільки один раз. У стовпчику з номером i виписано всі імпліканти функції f довжини $n - i$.

Процедура закінчується тоді, коли не можна виконати жодного склеювання.

Множини номерів разом з їх векторами індексів, які не позначено зірочками в підсумковій таблиці, визначають усі прості імпліканти функції f.

Покажемо, як можна знайти імпліканту довжини m, що відповідає такій парі: множина номерів P і вектор індексів $(i_1, i_2, \ldots, i_k)$, $k = n - m$. Випишемо будь-яку елементарну кон'юнкцію K, що накриває двійковий кортеж довжини n, номер якого входить у P. Викреслимо з елементарної кон'юнкції K змінні з номерами $n - \log_2 i_1$,

$n - \log_2 i_2$, ..., $n - \log_2 i_k$. Одержана елементарна кон'юнкція є шуканою імплікантою.

Застосувавши цей метод до рядків табл. 6.7, отримаємо шукану СДНФ функції f:

$$f(x, y, z, w) = \bar{x}\, y\, \bar{w} \vee \bar{y}\,\bar{z} \vee \bar{z}\,\bar{w} \vee \bar{y}\, w \vee x\bar{z}.$$

Одержана ДНФ не тупикова, а отже, не мінімальна, оскільки серед її імплікант є зайві. У наступному розділі розглянемо методи побудови всіх тупикових ДНФ функції f, виходячи з СДНФ, і пошуку серед них мінімальної ДНФ.

6.16. Задачі і вправи

1. За допомогою методу Квайна побудувати СДНФ функції f, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (1011)$;

(б) $\boldsymbol{a}_f = (10101001)$;

(в) $\boldsymbol{a}_f = (0000110000110011)$.

2. За допомогою методу Квайна побудувати СДНФ функції f, заданої формулою A:

(а) $A = ((\bar{y} \to (x \to \bar{z})) \mid ((((y \oplus z) \vee x) \mid ((\bar{z} \vee y) \to x)) \sim x))$;

(б) $A = ((x \vee y) \to (x \mid (y \wedge z))) \oplus ((y \sim z) \to x)$.

3. Методом Мак-Класкі побудувати СДНФ функції f, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (11000111)$;

(б) $\boldsymbol{a}_f = (10101100)$;

(в) $\boldsymbol{a}_f = (1000011100011110)$.

6.18. Методи побудови мінімальних ДНФ булевої функції

Із теореми 6.10 та її наслідку випливає, що для знаходження мінімальної ДНФ булевої функції f слід проаналізувати всі її тупикові ДНФ, а для одержання всіх тупикових ДНФ потрібно знайти й вилучити зі скороченої ДНФ усі можливі набори зайвих простих імплікант функції f.

Основний засіб розв'язання останньої задачі — побудова й аналіз так званої імплікантної таблиці функції f.

Імплікантна таблиця n-арної булевої функції f — це прямокутна таблиця з двома входами. Рядки таблиці позначають простими імплікантами K_i функції f, а стовпчики — двійковими кортежами σ_j довжини n (або їх номерами), на яких функція f набуває значення 1, $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, s$. Якщо імпліканта K_i накриває двійковий кортеж σ_j, то на перетині i-го рядка й j-го стовпчика таблиці ставлять зірочку. Всі інші клітинки таблиці залишаються порожніми.

Наприклад, для булевої функції $f(x, y, z, w)$ з прикладу 6.9 імплікантна таблиця має такий вигляд (табл. 6.8).

Таблиця 6.8

Проста імпліканта	**0**	**1**	**3**	**4**	**6**	**8**	**9**	**11**	**12**	**13**
$\bar{x}\, y\bar{w}$				*	*					
$\bar{y}\,\bar{z}$	*	*				*	*			
$\bar{z}\bar{w}$	*			*		*			*	
$\bar{y}\, w$		*	*				*	*		
$x\bar{z}$						*	*		*	*

Множина імплікант T накриває всі одиниці функції f, якщо в рядках імплікантної таблиці, що відповідають імплікантам множини T, кожен стовпчик містить принаймні одну зірочку. Називатимемо цю умову ***умовою покриття***.

Таким чином, деякий набір імплікант визначає ДНФ булевої функції f тоді й лише тоді, коли для відповідного набору рядків її імплікантної таблиці виконується умова покриття.

Тупиковим ДНФ відповідають набори рядків, з яких не можна вилучити жодного рядка, не порушуючи умови покриття. Назвемо такий набір рядків ***тупиковим покриттям*** імплікантної таблиці.

Процедура знаходження всіх тупикових покриттів імплікантної таблиці пов'язана, взагалі кажучи, зі значним перебором. Однак можна сформулювати деякі очевидні критерії, що дають змогу скоротити цей перебір.

1. Якщо в якомусь стовпчику імплікантної таблиці міститься тільки одна зірочка, то просту імпліканту, яка стоїть у відповідному рядку, будемо називати ***істотною***. Істотні імпліканти входять у всі тупикові ДНФ функції f, а отже, і в усі її мінімальні ДНФ. Не використовуючи ці імпліканти, неможливо накрити всі одиниці функції f.

Випишемо всі істотні прості імпліканти й викреслимо в імплікантній таблиці всі відповідні рядки та всі стовпчики для тих двійкових кортежів, що накриваються істотними імплікантами.

Наприклад, у табл. 6.8 істотними є імпліканти $\bar{x}y\bar{w}$, $\bar{y}\ w$ й xz. Після відповідних викреслень одержимо табл. 6.9.

Таблиця 6.9

Проста імпліканта	**0**
$y\bar{z}$	*
$\bar{z}\bar{w}$	*

2. Якщо в одержаній таблиці є два стовпчики, що містять зірочки в тих самих рядках, то один із цих стовпчиків можна викреслити, бо відповідні двійкові кортежі накриваються одними й тими самими імплікантами.

3. Якщо після викреслення стовпчиків у таблиці з'явилися рядки, що не містять жодної зірочки, то ці рядки (і відповідні їм імпліканти) викреслюють. В одержаній таблиці шляхом перебору всіх можливих варіантів знаходять усі тупикові покриття. До відповідних ДНФ додають усі істотні прості імпліканти функції f. Обчислюючи складність кожної з них, знайдемо всі мінімальні ДНФ.

Наприклад, у функції $f(x, y, z, w)$ з прикладу 6.9 існує дві тупикові ДНФ: $T_1 = \bar{x}y\bar{w} \vee \bar{y}\ w \vee x\bar{z} \vee \bar{y}\bar{z}$ і $T_2 = \bar{x}y\bar{w} \vee \bar{y}\ w \vee x\bar{z} \vee \bar{z}\bar{w}$.

Складності цих формул однакові, отже обидві вони є мінімальними ДНФ функції f.

Зазначимо, що перебір різних множин рядків для пошуку всіх тупикових покриттів можна практично виконати лише для відносно невеликих імплікантних таблиць. Можна також використовувати перебір, коли стоїть задача знайти не обов'язково мінімальну ДНФ, а достатньо одержати першу-ліпшу ДНФ заданої функції, простішу від скороченої ДНФ.

Зменшити об'єм обчислень при знаходженні всіх тупикових покриттів імплікантної таблиці можна за допомогою формальної процедури, запропонованої С. Петріком.

Метод Петріка побудови всіх тупикових ДНФ булевої функції

Кожній простій імпліканті булевої функції f, тобто кожному рядку імплікантної таблиці функції f, поставимо у відповідність нову булеву

змінну. Наприклад, для функції $f(x, y, z)$ із прикладу 6.6 одержимо таку імплікантну таблицю (табл. 6.10).

Таблиця 6.10

Нова змінна	**Проста імпліканта**	**1**	**2**	**3**	**4**	**5**	**6**
a	$\bar{x}z$	*		*			
b	$\bar{y}z$	*				*	
c	xy		*	*			
d	$y\bar{z}$		*				*
e	xy				*	*	
h	$x\bar{z}$				*		*

Кожній множині P простих імплікант функції f відповідає набір значень нових змінних: якщо якась імпліканта входить у множину P, то відповідна змінна дорівнює 1, якщо не входить — 0.

Для кожного стовпчика σ_j імплікантної таблиці утворимо диз'юнкцію тих булевих змінних, відповідні рядки яких містять зірочки у стовпчику σ_j. Утворена диз'юнкція $a_1 \vee a_2 \vee \ldots \vee a_k$ дорівнює 1 тоді й тільки тоді, коли принаймні одна з відповідних імплікант накриває двійковий кортеж σ_j.

Кон'юнкція всіх утворених таким чином диз'юнкцій дорівнює 1 тоді й тільки тоді, коли кожна з диз'юнкцій дорівнює 1, тобто коли кожен із двійкових кортежів з множини N_f буде накритий якоюсь простою імплікантою.

Побудуємо зазначену кон'юнкцію для табл. 6.10. Для її стовпчиків одержимо шість диз'юнкцій: $a \vee b$, $c \vee d$, $a \vee c$, $e \vee h$, $b \vee e$, $d \vee h$. Кон'юнкція всіх цих диз'юнкцій дає таку формулу:

$$F = (a \vee b)(c \vee d)(a \vee c)(e \vee h)(b \vee e)(d \vee h).$$

Ще раз підкреслимо, що коли на якомусь наборі значень $\alpha \in B^6$ змінних a, b, c, d, e, h формула F набуває значення 1, то набір імплікант, яким в α відповідають 1, накриває всі одиниці функції f.

Перетворимо формулу F до диз'юнктивної нормальної форми. Для цього за законом дистрибутивності розкриємо в ній дужки. При цьому будемо використовувати правила поглинання (6.4): $x(x \vee y) = x$, $x \vee xy = x$, а також ідемпотентність кон'юнкції: $xx = x$.

$F = (ac \vee bc \vee ad \vee bd)(a \vee c)(be \vee bh \vee ee \vee eh)(d \vee h) =$
$= (ac \vee abc \vee ad \vee abd \vee ac \vee bc \vee acd \vee bcd) \wedge (bde \vee bdh \vee de \vee$
$\vee deh \vee beh \vee bh \vee eh \vee eh) = (ac \vee bc \vee ad)(de \vee eh \vee bh) =$
$= acde \vee bcde \vee ade \vee aceh \vee bceh \vee bch \vee ade \vee adeh \vee abdh =$
$= ade \vee bch \vee bcde \vee aceh \vee abdh.$

Останній вираз набуває значення 1 тоді й тільки тоді, коли дорівнює 1 принаймні одна з його елементарних кон'юнкцій. У свою чергу, елементарна кон'юнкція набуває значення 1, коли всі її операнди дорівнюють 1. Отже, кожна з цих елементарних кон'юнкцій визначає певне тупикове покриття імплікантної таблиці. Таким чином, функція $f(x, y, z)$ має п'ять тупикових ДНФ:

$T_1 = \bar{x}z \vee y\bar{z} \vee x\bar{y}$, $\quad T_4 = \bar{x}z \vee \bar{x}y \vee x\bar{y} \vee x\bar{z}$,
$T_2 = \bar{y}z \vee \bar{x}y \vee x\bar{z}$, $\quad T_5 = \bar{x}z \vee \bar{y}z \vee y\bar{z} \vee x\bar{z}$.
$T_3 = \bar{y}z \vee \bar{x}y \vee y\bar{z} \vee x\bar{y}$,

Три з них (T_5, T_2 і T_1) було побудовано у прикладі 6.7 шляхом вилучення зайвих імплікант зі скороченої ДНФ функції f.

Одержавши описаним методом усі тупикові ДНФ й обчисливши складність кожної з них, визначимо всі мінімальні ДНФ заданої функції f.

Зі зростанням числа змінних кількість можливих тупикових ДНФ функції f для абсолютної більшості булевих функцій зростає надзвичайно швидко, тому описані методи мінімізації застосовують лише для функцій від порівняно невеликої кількості змінних.

Для булевих функцій від великої кількості змінних розробляють і застосовують різноманітні наближені методи мінімізації, які, взагалі кажучи, не гарантують одержання мінімальної ДНФ заданої функції, але за невелику кількість кроків і з витратою значно менших ресурсів дають змогу будувати ДНФ, що мало відрізняються за складністю від мінімальної.

6.17. Задачі і вправи

1. Побудувати всі тупикові й усі мінімальні ДНФ булевої функції f, заданої вектором значень $\boldsymbol{a}_f$:

(а) $\boldsymbol{a}_f = (01101110)$;
(б) $\boldsymbol{a}_f = (01011100)$;
(в) $\boldsymbol{a}_f = (1110011001010101)$.

2. Виділити зі скороченої ДНФ усі істотні імпліканти:

(а) $\bar{x}z \vee y\bar{z} \vee x\bar{y} \vee \bar{x}\bar{y}w \vee \bar{y}zw \vee xzw$;
(б) $xyz \vee x\bar{y}z \vee \bar{z}w \vee \bar{y}w \vee \bar{x}yz \vee \bar{x}w$;
(в) $xw \vee \bar{y}z \vee y\bar{w} \vee \bar{x}yzw$.

3. З'ясувати, чи є тупиковою, найкоротшою чи мінімальною така ДНФ:
(а) $xy \vee \bar{y}$; (б) $\bar{z}\bar{w} \vee x\bar{y}w \vee \bar{x}\bar{y}z$; (в) $\bar{x}z \vee xy \vee \bar{x}\bar{y}w \vee xzw$.

4. Визначити кількість ТДНФ і МДНФ булевої функції, заданої вектором значень (0111111111111110).

6.19. Інші методи мінімізації ДНФ булевих функцій

В описаному вище методі Квайна–Мак-Класкі передбачається, що побудова скороченої ДНФ починається з досконалої ДНФ. Однак часто виникає ситуація, коли булеву функцію зображено довільною ДНФ. У цьому разі доцільно використовувати метод мінімізації, запропонований А. Блейком [8].

Метод Блейка

Неважко переконатися, що в алгебрі логіки виконується співвідношення, яке називають ***формулою узагальненого склеювання*** (за змінною x)

$$Ax \vee B\bar{x} = Ax \vee B\bar{x} \vee AB. \qquad (6.16)$$

Справді, запишемо такі рівносильні перетворення:

$$Ax \vee B\bar{x} = Ax \vee ABx \vee B\bar{x} \vee AB\bar{x} = Ax \vee B\bar{x} \vee AB(x \vee \bar{x}) = Ax \vee B\bar{x} \vee AB.$$

Оскільки множина всіх можливих елементарних кон'юнкцій (й імплікант) булевої функції f скінченна, то, послідовно застосовуючи до членів ДНФ функції f формулу узагальненого склеювання, через певну кількість перетворень у формулі перестануть з'являтися нові елементарні кон'юнкції.

Результат Блейка полягає в тому, що коли в довільній ДНФ функції f виконати всі можливі узагальнені склеювання й усі елементарні поглинання, то в результаті одержимо СДНФ функції f [8].

Приклад 6.10. Розглянемо процедуру побудови СДНФ функції f, заданої ДНФ $F = \bar{x}\bar{y} \vee \bar{x}\bar{z} \vee x\bar{y}z \vee xyz$, за методом Блейка. Для зручності для кожного склеювання формули F після рівності в дужках будемо записувати ту змінну, за якою виконано узагальнене склеювання:

$x\bar{y}z \vee xyz = x\bar{y}z \vee xyz \vee xz \quad (y)$,
$\bar{x}\bar{y} \vee x\bar{y}z = \bar{x}\bar{y} \vee x\bar{y}z \vee \bar{y}\bar{y}z = \bar{x}\bar{y} \vee x\bar{y}z \vee \bar{y}z \quad (x)$,
$\bar{x}\bar{y} \vee xyz = \bar{x}\bar{y} \vee xyz \vee \bar{y}z = \bar{x}\bar{y} \vee xyz \quad (x)$,
$\bar{x}\bar{y} \vee xyz = \bar{x}\bar{y} \vee xyz \vee \bar{x}z = \bar{x}\bar{y} \vee xyz \quad (y)$,
$\bar{x}\bar{z} \vee x\bar{y}z = \bar{x}\bar{z} \vee x\bar{y}z \vee \bar{z}\bar{y}z = \bar{x}\bar{z} \vee x\bar{y}z \quad (x)$,

$\bar{x}\bar{z} \vee x\bar{y}z = \bar{x}\bar{z} \vee x\bar{y}z \vee \bar{x}\,\bar{y} = \bar{x}\bar{z} \vee x\bar{y}z$ (z),
$\bar{x}\bar{z} \vee xyz = \bar{x}\bar{z} \vee xyz \vee \bar{z}yz = \bar{x}\bar{z} \vee xyz$ (x),
$\bar{x}\bar{z} \vee xyz = \bar{x}\bar{z} \vee xyz \vee \bar{x}y = \bar{x}\bar{z} \vee xyz$ (z).

Після всіх узагальнених склеювань членів формули F одержано лише дві нові імпліканти xz і $\bar{y}z$. Продовжимо процес склеювання з цими імплікантами:

$\bar{x}\bar{y} \vee xz = \bar{x}\bar{y} \vee xz \vee \bar{y}z$ (x),
$\bar{x}\bar{z} \vee xz = \bar{x}\bar{z} \vee xz \vee \bar{z} = \bar{x}\bar{z} \vee xz$ (x),
$\bar{x}\bar{z} \vee xz = \bar{x}\bar{z} \vee xz \vee \bar{x} = \bar{x}\bar{z} \vee xz$ (z),
$\bar{x}\bar{z} \vee \bar{y}z = \bar{x}\bar{z} \vee \bar{y}z \vee \bar{x}\bar{y}(z)$.

Процес стабілізувався; виконавши всі можливі елементарні поглинання, одержимо шукану СДНФ функції f:

$$\bar{x}\bar{y} \vee xz \vee \bar{y}z \vee \bar{x}\bar{z}$$

Усі описані вище й інші подібні до них методи мінімізації ДНФ булевої функції називають *аналітичними*. На практиці для мінімізації ДНФ булевої функції, яка залежить від малого числа змінних, часто застосовують зручніші й наочніші, так звані *графічні*, або *табличні*, методи.

Карти Карно, або діаграми Вейча

Основна ідея графічних методів полягає в зображенні таблиць істинності булевих функцій, для яких потрібно знайти мінімальні ДНФ, у спеціальному вигляді, а саме у вигляді так званих карт Карно (або діаграм Вейча).

Карти Карно для функцій від двох, трьох і чотирьох змінних зображено на рис. 6.13 (*а* – *в*).

а) $n = 2$

$x \setminus y$	**0**	**1**
0	f_0	f_1
1	f_2	f_3

б) $n = 3$

$xy \setminus z$	**0**	**1**
00	f_0	f_1
01	f_2	f_3
11	f_6	f_7
10	f_4	f_5

в) $n = 4$

$xy \setminus zw$	**00**	**01**	**11**	**10**
00	f_0	f_1	f_3	f_3
01	f_4	f_5	f_7	f_6
11	f_{12}	f_{13}	f_{15}	f_{14}
10	f_8	f_9	f_{11}	f_{10}

Рис. 6.13

Через f_j у картах Карно позначають значення заданої булевої функції на кортежі ***a*** (відповідної довжини) з номером j. Відповідний кортеж одержимо, якщо випишемо спочатку значення змінних із відповідного рядка, а потім — із відповідного стовпчика. Наприклад, для булевої функції від трьох змінних $f_4 = f(1, 0, 0)$, $f_2 = f(0, 1, 0)$, а для булевої функції від чотирьох змінних $f_9 = f(1, 0, 0, 1)$, $f_3 = f(0, 0, 1, 1)$.

Звернімо увагу на те, що двійкові кортежі значень змінних і відповідні їм значення функції розміщено в таблицях не у звичному порядку. Із цього спеціального порядку випливає важлива особливість карти Карно: будь-яка елементарна кон'юнкція змінних накриває в ній *прямокутник* зі сторонами, що є степенями 2.

Пояснимо проголошену тезу на прикладах.

Для карти Карно булевої функції від двох змінних елементарна кон'юнкція $\bar{x}y$ накриває прямокутник цієї карти, що складається з єдиного елемента f_1, тобто прямокутник розміром 1×1; елементарна кон'юнкція $\bar{x}$ накриває прямокутник розміром 1×2, що складається з двох елементів f_0, f_1. Аналогічно, $\bar{x}\bar{y}$ накриває 1×1-прямокутник f_0, а $\bar{y}$ та x — відповідно прямокутники f_0, f_2 (розміром 2×1) і прямокутник f_2, f_3 (розміром 1×2).

Розглядаючи карту Карно для булевої функції від трьох змінних, узагальнимо поняття прямокутника, склеївши (з'єднавши) таблицю 6.13, *б* за її верхнім і нижнім краями. Отже, клітинки пар f_0 і f_4 та f_1 і f_5 стануть сусідніми.

Тоді для карти Карно булевої функції від трьох змінних імпліканта $\bar{x}yz$ накриває 1×1-прямокутник f_3, імпліканта $y\,\bar{z}$ — 2×1-прямокутник f_2, f_6, імпліканта $\bar{z}$ — 4×1-прямокутник f_0, f_2, f_6, f_4, імпліканта $\bar{y}$ — 2×2-прямокутник f_4, f_5, f_0, f_1 і т. д.

У карті Карно для булевої функції від чотирьох змінних окрім склеювання нижнього і верхнього країв таблиці 6.13, *в* потрібно також склеїти лівий і правий її краї. Тоді імпліканта $x\bar{z}$ накриває 2×2-прямокутник f_{12}, f_{13}, f_8, f_9; імпліканта zw — 4×1-прямокутник f_2, f_6, f_{14}, f_{10}; імпліканта $\bar{x}\,\bar{w}$ — 2×2-прямокутник f_2, f_0, f_6, f_4; імпліканта $\bar{y}\,\bar{w}$ — 2×2-прямокутник f_2, f_0, f_{10}, f_8. Для імплікант довжини 1 маємо: $\bar{x}$ накриває 2×4-прямокутник із двох перших рядків таблиці 6.13, *в*, $\bar{z}$ — 4×2-прямокутник із двох перших стовпчиків, $\bar{y}$ — 2×4-прямокутник, що складається з першого й останнього рядків.

Графічний метод побудови мінімальної ДНФ булевої функції f складається з таких кроків.

1. Сформувати карту Карно для функції f.

2. Знайти покриття всіх одиниць функції f прямокутниками максимальних розмірів так, щоб їх кількість була найменшою. При цьому може трапитися, що деякі одиниці буде накрито кількома різними прямокутниками.

3. Диз'юнкція імплікант, що відповідають прямокутникам покриття, є шуканою мінімальною ДНФ.

Продемонструємо застосування графічного методу мінімізації ДНФ на прикладах.

Приклад 6.11. Розглянемо чотири булеві функції від трьох змінних $f_i(x, y, z)$, $i = 1, 2, 3, 4$:

$N_{f_1} = \{(0, 0, 0), (0, 0, 1), (0, 1, 1), (1, 1, 0), (1, 1, 1)\}$,

$N_{f_2} = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (1, 0, 1), (1, 1, 0)\}$,

$N_{f_3} = \{(0, 0, 0), (0, 1, 1), (1, 1, 0)\}$,

$N_{f_4} = \{(0, 0, 0), (0, 1, 1), (1, 0, 0), (1, 1, 1)\}$.

Карти Карно цих функцій подано на рис. 6.14.

Зрозуміло, що оскільки булева функція може мати кілька мінімальних ДНФ, можуть бути різні варіанти оптимального покриття її карти Карно. Для кожної з карт зазначимо один із можливих варіантів покриття набором прямокутників:

$\{(f_0, f_1), (f_1, f_3), (f_6, f_7)\}$ — для функції $f_1(x, y, z)$,

$\{(f_0, f_1), (f_0, f_2, f_6, f_4), (f_4, f_5)\}$ — для функції $f_2(x, y, z)$,

$\{(f_0), (f_3), (f_6)\}$ — для функції $f_3(x, y, z)$,

$\{(f_0, f_4), (f_3, f_7)\}$ — для функції $f_4(x, y, z)$.

f_1

xy\z	0	1
00	1	1
01	0	1
11	1	1
10	0	0

f_2

xy\z	0	1
00	1	1
01	1	0
11	1	0
10	1	1

f_3

xy\z	0	1
00	1	0
01	0	1
11	1	0
10	0	0

f_4

xy\z	0	1
00	1	0
01	0	1
11	0	1
10	1	0

Рис. 6.14

Відповідні мінімальні ДНФ дорівнюють:

$F_1 = \bar{x}\bar{y} \vee \bar{x}z \vee xy$, $\quad F_3 = \bar{x}\bar{y}\bar{z} \vee \bar{x}y\,z \vee xy\bar{z}$,

$F_2 = \bar{y} \vee \bar{z}$, $\quad F_4 = yz \vee \bar{y}\bar{z}$.

У карті Карно функції $f_3(x, y, z)$ жодну з одиниць не накриває прямокутник розміром більшим ніж 1×1, а тому мінімальна ДНФ збігається з ДДНФ функції $f_3(x, y, z)$.

Приклад 6.12. Чотири булеві функції $f_i(x, y, z, w)$, $i = 1, 2, 3, 4$, від чотирьох змінних задано множинами N_{f_i} двійкових кортежів (для зручності замість двійкових кортежів пишемо їх номери):

$N_{f_1} = \{1, 5, 6, 9, 11, 13, 14\}$,

$N_{f_2} = \{0, 1, 2, 5, 7, 8, 9, 10, 13\}$,

$N_{f_3} = \{0, 1, 2, 4, 6, 7, 8, 9\}$,

$N_{f_4} = \{0, 7, 10, 11, 12\}$.

Карти Карно цих функцій подано на рис. 6.15.

Можливими покриттями для кожної з карт можуть бути, наприклад, такі набори прямокутників:

$\{(f_1, f_5, f_{13}, f_9), (f_9, f_{11}), (f_6, f_{14})\}$ — для функції $f_1(x, y, z, w)$,

f_1

$xy \backslash zw$	00	01	11	10
00	0	1	0	0
01	0	1	0	1
11	0	1	0	1
10	0	1	1	0

f_2

$xy \backslash zw$	00	01	11	10
00	1	1	0	1
01	0	1	1	0
11	0	1	0	0
10	1	1	0	1

f_3

$xy \backslash zw$	00	01	11	10
00	1	1	0	1
01	1	0	1	1
11	0	0	0	0
10	1	1	0	0

f_4

$xy \backslash zw$	00	01	11	10
00	1	0	0	0
01	0	0	1	0
11	1	0	0	0
10	0	0	1	1

Рис. 6.15

$\{(f_0, f_2, f_8, f_{10}), (f_1, f_5, f_{13}, f_9), (f_5, f_7)\}$ — для функції $f_2(x, y, z, w)$,

$\{(f_0, f_1, f_8, f_9), (f_0, f_4, f_2, f_6), (f_7, f_6)\}$ — для функції $f_3(x, y, z, w)$,

$\{(f_0), (f_7), (f_{12}), (f_{11}, f_{10})\}$ — для функції $f_4(x, y, z, w)$.

Відповідні мінімальні ДНФ функцій f_1, f_2, f_3, f_4 дорівнюють:

$F_1 = \bar{z}w \vee yz\bar{w} \vee x\bar{y}w,$ $\qquad F_3 = \bar{x}\bar{w} \vee \bar{y}\bar{z} \vee \bar{x}yz,$

$F_2 = \bar{y}\bar{w} \vee \bar{z}w \vee \bar{x}yw,$ $\qquad F_4 = x\bar{y}z \vee \bar{x}\bar{y}\bar{z}\bar{w} \vee \bar{x}yzw \vee xy\bar{z}\bar{w}.$

Карти Карно булевих функцій арності, більшої 4, не можна побудувати за безпосередньою аналогією з картами для $n = 2, 3, 4$; вони істотно ускладнюються. Потрібно узагальнити відношення сусідства клітин і поняття прямокутників у таких таблицях. Для пошуку прямокутників максимального розміру, що накривають усі одиниці заданої функції, потрібна значно більша просторова уява й увага. Це призводить до того, що карти Карно використовують для мінімізації ДНФ булевих функцій арності не більшої ніж 8.

Іноді карти Карно булевих функцій для $n = 5, 6, 7, 8$ зображають не прямокутними таблицями на площині, а у вигляді прямокутних паралелепіпедів у просторі. Тоді роль накривальних прямокутників виконують прямокутні паралелепіпеди.

Для булевих функцій від більшого числа змінних у загальному випадку процедура побудови мінімальних ДНФ за допомогою графічного методу стає занадто громіздкою та незручною. У цих ви-падках застосовують аналітичні методи мінімізації типу методів Квайна, Мак-Класкі, Блейка та ін. [8].

6.19. Задачі і вправи

1. За допомогою методу Блейка побудувати СДНФ функції f за заданою її ДНФ D:

(а) $D = \bar{x}\bar{y} \vee x\bar{y}z \vee \bar{x}z$;

(б) $D = \bar{x}\bar{y}z \vee x\bar{y}z \vee \bar{x}\bar{z}$;

(в) $D = xy \vee \bar{x}z \vee x\bar{y}zw \vee \bar{x}\bar{y}zw$.

2. Побудувати СДНФ за заданою КНФ:

(а) $(x \vee y \vee \bar{z})(\bar{x} \vee y \vee z)(\bar{y} \vee z)$;

(б) $(x \vee \bar{y} \vee \bar{z})(\bar{x} \vee w)(y \vee z \vee \bar{w})$.

3. За допомогою карти Карно визначити всі мінімальні ДНФ булевої функції f, заданої вектором значень a_f:

(а) $a_f = (01010011)$;

(б) $a_f = (10001010)$;

(в) $a_f = (0110101111011110)$.

РОЗДІЛ 7

ТЕОРІЯ АВТОМАТІВ

Однією з найпопулярніших і найпоширеніших математичних моделей, які активно й успішно використовувались і використовуються у теорії та практиці сучасної комп'ютерної науки (інформатики), є так званий скінченний автомат. Варто назвати лише деякі найважливіші *приклади застосування* цієї моделі.

1. Опис функціонування та логічного проектування найрізноманітніших схем і пристроїв дискретної дії. Зокрема, скінченний автомат широко використовують для розробки й аналізу схем і пристроїв обчислювальної техніки, як модель для інтерпретування мікропрограм виконання команд ЕОМ та ін.

2. Скінченний автомат — зручний і ефективний засіб для опису та створення важливого компонента стандартного компілятора мови програмування, який називають *лексичним аналізатором*. Цей програмний модуль відповідає за розбивку вхідного тексту на логічні одиниці: ідентифікатори, ключові (резервовані) слова, константи, спеціальні знаки тощо.

3. Ця модель є складовою частиною програмного забезпечення, призначеного для перегляду великих текстових масивів даних (наприклад, набору Web-сторінок) з метою пошуку в них потрібної інформації. Ключем для такого пошуку служать так звані "ключові слова", або певні послідовності символів ("образи").

4. Модель скінченного автомата використовують у програмному забезпеченні для розробки й аналізу поведінки різноманітних систем, що можуть перебувати у скінченному числі відмінних один від одного станів. Прикладами таких систем можуть бути протокол безпечного (захищеного) обміну інформацією, протокол, що організує та контролює процедури електронної торгівлі та інших фінансових операцій у комп'ютерних мережах, тощо.

7.1. Поняття скінченного автомата. Методи задання автоматів

Скінченним автоматом (у подальшому — просто ***автоматом*** або *СА*) називається система $A = (X, Y, U, \delta, \lambda)$, в якій $X = \{x_1, x_2, \ldots, x_m\}$ та $Y = \{y_1, y_2, \ldots, y_n\}$ — скінченні ***множини*** (*алфавіти*) відповідно ***вхідних і вихідних сигналів***, $U = \{a_1, a_2, \ldots, a_s\}$ — ***множина внутрішніх станів*** автомата. Функції δ й λ описують алгоритм функціонування (поведінку) автомата A. Функція $\delta{:}U \times X \to U$ називається ***функцією переходів***, а $\lambda{:}U \times X \to Y$ — ***функцією виходів*** автомата A.

Необхідно підкреслити одну особливість моделі СА, яка випливає з її подальшої фізичної інтерпретації. Вважаємо, що автомат A функціонує в дискретному часі, тобто час функціонування автомата розбито на відрізки (однакової чи різної довжини) — *такти*. Межі тактів t називають моментами абстрактного дискретного ***автоматного часу*** й нумерують цілими числами, починаючи з нуля. Значення вхідних і вихідних сигналів та значення станів автомата можуть змінюватися тільки в моменти автоматного часу $t = 1, 2, \ldots$.

Якщо позначимо через $x(t)\in X$, $y(t)\in Y$, $a(t)\in U$ значення вхідного та вихідного сигналів і стану автомата в момент автоматного часу t, то роботу автомата A можна описати такими співвідношеннями:

$$a(t+1) = \delta\big(a(t), x(t)\big), \quad y(t) = \lambda\big(a(t), x(t)\big). \tag{7.1}$$

Співвідношення (7.1) називають ***канонічними рівняннями*** автомата A.

Перше з канонічних рівнянь можна прочитати так: стан автомата A в будь-який момент автоматного часу $t + 1$ визначається сигналом, поданим на вхід автомата, і станом автомата A в попередній момент автоматного часу. При цьому кажуть, що автомат A переходить зі стану $a(t)$ у стан $a(t + 1)$.

У математичній моделі автомата можна вважати, що цей перехід відбувається миттєво, або стрибкоподібно. При фізичній інтерпретації розглядуваної моделі слід брати до уваги, що зазначений перехід відбувається поступово та потребує певного фізичного часу. Однак вважають, що цей час завжди менший від тривалості такту введеного дискретного часу; отже, від моменту t до моменту $t + 1$ весь перехідний процес завершується.

Крім зміни станів результатом роботи автомата є також видача вихідних сигналів за законом, визначеним другим канонічним рівнянням автомата.

Якщо в автоматі A виділено стан, у якому автомат A перебуває в момент автоматного часу $t = 0$, то цей стан називають ***початковим*** (зазвичай початковим станом уважають a_1), а автомат A називають ***ініціальним*** і позначають A/a_1. Надалі не будемо вказувати явно залежність змінних і результатів функцій переходів і виходів від автоматного часу t, крім тих випадків, коли це потрібно.

Для розв'язання різних задач теорії автоматів зручно використовувати різні ***способи (методи) задання автоматів***. Опишемо три найпоширеніші з них. В усіх трьох методах істотно використано те, що функції δ і λ автомата A скінченно-визначені, тобто мають скінченні області визначення.

Табличний спосіб. Функції δ й λ можна задавати за допомогою двох таблиць, які називають відповідно ***таблицею переходів*** і ***таблицею виходів*** автомата A. Загальна структура обох таблиць однакова: рядки таблиць позначено вхідними сигналами $x_1, x_2, \ldots, x_m$, а стовпчики — станами $a_1, a_2, \ldots, a_s$. На перетині i-го рядка та j-го стовпчика в таблиці переходів записують стан $\delta(a_j, x_i)$, а в таблиці виходів — вихідний сигнал $\lambda(a_j, x_i)$. Іноді для задання автомата використовують одну ***суміщену таблицю*** *переходів/виходів*, у якій на перетині i-го рядка та j-го стовпчика записують відповідну пару a_k/y_l, де $a_k = \delta(a_j, x_i)$ й $y_l = \lambda(a_j, x_i)$.

Графічний спосіб. Автомат можна задати за допомогою орієнтованого мультиграфа, який називають ***графом***, або ***діаграмою автомата*** (*автоматним графом*, *автоматною діаграмою*). Вершини графа позначають символами станів автомата A. Якщо $\delta(a_i, x_k) = a_j$ і $\lambda(a_i, x_k) = y_l$, то в графі автомата проводять орієнтовану дугу (або стрілку) з вершини a_i у вершину a_j і позначають цю дугу символами x_k/y_l. Задання автомата за допомогою графа особливо наочне, якщо кількість його станів порівняно невелика.

Матричний спосіб. Функцію переходів автомата A можна задати за допомогою системи m $s{\times}s$-***матриць переходів*** $M(x_1), M(x_2), \ldots$ $\ldots, M(x_m)$. Якщо позначити через $m_{ij}(x_k)$ елемент i-го рядка та j-го стовпчика матриці $M(x_k)$, то $m_{ij}(x_k) = 1$, коли $a_j = \delta(a_i, x_k)$, і $m_{ij}(x_k) = 0$ в іншому разі.

Аналогічно означають m $s{\times}n$-***матриць виходів*** $W(x_1), W(x_2), \ldots, W(x_m)$ для функції виходів λ. Якщо $w_{ij}(x_k)$ — елемент i-го рядка та

j-го стовпчика матриці $W(x_k)$, то $w_{ij}(x_k) = 1$, коли $y_j = \lambda(a_i, x_k)$, і $w_{ij}(x_k) = 0$ в іншому разі.

Зрозуміло, що досить легко можна перейти від одного способу задання до будь-якого іншого.

Автомат A називають ***детермінованим***, якщо функції δ і λ всюди визначені, і ***недетермінованим***, якщо δ і λ — всюди визначені відповідності між $U \times X$ і U та $U \times X$ і Y відповідно, але вони не задовольняють умову функціональності. Автомат A називають ***частковим***, якщо функції δ і λ не є всюди визначеними.

Приклад 7.1. Розглянемо автомат D, який регулює дорожній рух на перехресті вулиць В і П.

В автомат дорожнього руху D з періодом 1 хв надходить тактовий сигнал Г генератора синхроімпульсів, що послідовно перемикає сигнали світлофора С, дозволяючи транспорту рух відповідно вулицями В і П (рис. 7.1, *а*). Крім світлофора є також кнопка виклику К, за допомогою якої пішохід може надіслати автомату запит З на призупинення руху на перехресті. При надходженні запиту З та після завершення поточного інтервалу часу в 1 хв автомат перериває генерування послідовності сигналів В і П на 2 хв, сигналом Д запалює транспарант, що дозволяє перехід пішоходам, а через 2 хв формує сигнал скидання СС, повертаючи автомат до відновлення формування послідовності сигналів В та П. Отже, автомат D виробляє вихідні сигнали В, П, Д і СС відповідно дозволу руху транспорту по вулицях В і П, дозвіл переходу пішоходам і скидання кнопки виклику. Роботу автомата D ілюструє часова діаграма на рис. 7.1, *б*.

а *б*

Рис. 7.1

Таблиці переходів і виходів автомата дорожнього руху подано на рис. 7.2. Тут використано такі позначення: $x_1 = \text{Г}$, $x_2 = \text{З} \,\&\, \text{Г}$, $y_1 = \text{В}$, $y_2 = \text{П}$, $y_3 = \text{Д}$, $y_4 = \text{Д} \,\&\, \text{СС}$.

δ	a_1	a_2	a_3	a_4	a_5	a_6
x_1	a_2	a_1	a_4	a_2	a_6	a_1
x_2	a_3	a_5	a_4	a_2	a_6	a_1

λ	a_1	a_2	a_3	a_4	a_5	a_6
x_1	y_2	y_1	y_4	y_2	y_4	y_1
x_2	y_3	y_3	y_4	y_2	y_4	y_1

Рис. 7.2

Граф автомата дорожнього руху подано на рис. 7.3.

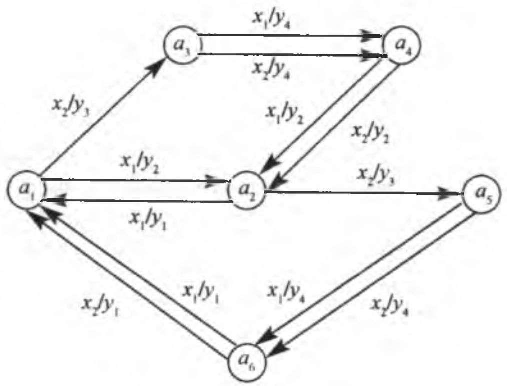


Рис. 7.3

І нарешті, розглянемо матричний спосіб задання функцій переходів і виходів автомата дорожнього руху D. Його матриці переходів і матриці виходів мають такий вигляд:

$$M(x_1) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad M(x_2) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$W(x_1) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad W(x_2) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Приклад 7.2. Побудуємо автомат S, який описує алгоритм функціонування послідовного двійкового суматора. На вхід автомата надходять пари розрядів двійкових чисел, що додаються, починаючи з молодших розрядів. На виході автомата має з'являтися розряд результату додавання. Потрібно також фіксувати й враховувати ситуацію наявності чи відсутності перенесення в наступний розряд. Ці дві ситуації відображають два стани автомата S: початковому стану a_1 відповідає ситуація "немає перенесення", а стану a_2 — ситуація "є перенесення".

Вхідні сигнали $x_0 = (0, 0)$, $x_1 = (0, 1)$, $x_2 = (1, 0)$ і $x_3 = (1, 1)$ задають чотири можливі комбінації розрядів, що додаються.

Граф автомата S подано на рис. 7.4.

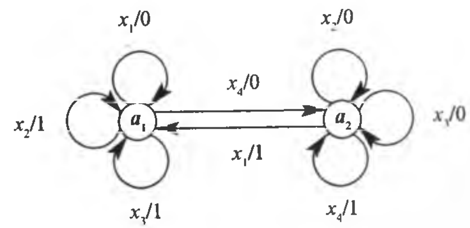


Рис. 7.4

Приклад 7.3. Розглянемо автомат Z, що виконує три операції зсуву дворозрядного двійкового коду на один розряд ліворуч або праворуч. Чотири стани автомата відповідають чотирьом можливим варіантам кодів: $a_0 = (0, 0)$, $a_1 = (0, 1)$, $a_2 = (1, 0)$ і $a_3 = (1, 1)$. Розряд, який у результаті зсуву виходить за межі розрядної сітки, вважаємо вихідним сигналом автомата Z.

Три вхідні сигнали визначають характер операції зсуву:

x_1 — код зсувається праворуч; у ліву позицію, звільнену в результаті зсуву, записується той розряд, який був у цій позиції до зсуву (знак "розмножується");

x_2 — код зсувається ліворуч; у праву звільнену позицію записується 0;

x_3 — код зсувається ліворуч; у праву звільнену позицію записується 1.

Граф автомата Z подано на рис. 7.5.

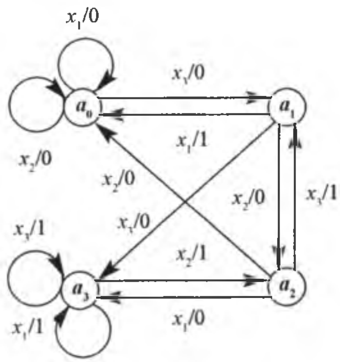


Рис. 7.5

7.1. Задачі і вправи

1. Побудувати таблиці переходів і виходів, граф, матриці переходів і виходів автомата Z, що реалізує три операції зсуву трирозрядного двійкового коду на одну позицію. Вхідні сигнали визначають характер операції зсуву: x_1 — логічний зсув ліворуч (праворуч), x_2 — арифметичний зсув праворуч, x_3 — циклічний зсув праворуч (ліворуч). Вихідний сигнал — це сигнал, що виходить за межі розрядної сітки. Стани відповідають восьми можливим варіантам кодів.

2. Побудувати таблиці переходів і виходів, граф, матриці переходів і виходів автомата P, що за допомогою своїх станів запам'ятовує три останні символи (двійкові розряди), які було подано на його вхід. Вихідним сигналом є останній символ, що "забувається".

3. Побудувати таблиці переходів і виходів, граф, матриці переходів і виходів автомата A, який переходить у стан T, якщо на його вхід було подано послідовність символів, що відповідає ідентифікатору якоїсь мови програмування, а не то переходить у стан H. Вхідні сигнали автомата A: x_1 — літера, x_2 — цифра, x_3 — будь-який інший символ.

4. Побудувати таблиці переходів і виходів, граф, матриці переходів і виходів автомата B, який переходить у стан T, якщо на його вхід було подано послідовність символів, що відповідає числовій константі — дійсному числу зі знаком і фіксованою крапкою, і переходить у стан H — в іншому разі. Вхідні сигнали автомата B: x_1 — знак, x_2 — цифра, x_3 — крапка, x_4 — будь-який інший символ.

5. Побудувати автомат затримки, тобто такий, вихідний сигнал якого в момент часу $t + 1$ дорівнює вхідному сигналу в момент часу t, $X = Y = \{0, 1\}$.

6. Побудувати автомат для $X = Y = \{0, 1\}$, на виході якого з'являється сигнал 1 тоді й лише тоді, коли чотири останні вхідні сигнали — 0110.

7. Побудувати "генератор парності", тобто автомат, що функціонує в алфавітах $X = Y = \{0, 1\}$ і реалізує такий алгоритм. На його вхід надходять слова довжини 3, розділені якимось символом a, $a \in X$. На виході автомат має повторити вхідну трійку символів, замінивши розділовий символ a на 1 тоді й лише тоді, коли кількість одиниць у даній трійці парна.

8. Побудувати автомат, що порівнює два додатні двійкові числа з однаковою розрядністю. Пари розрядів порівнюваних чисел подаються на вхід автомата, починаючи зі старших розрядів. Вихідним словом є більше з даних чисел.

9. Побудувати автомат, що перевіряє відповідність лівих і правих дужок у вхідному слові, яке задає певний алгебричний вираз. Найбільша дозволена вкладеність дужок — n. Вхідні сигнали автомата: x_1 — ліва дужка, x_2 — права дужка, x_3 — будь-який інший символ, x_4 — символ, що позначає кінець виразу.

10. Побудувати автомат, що керує роботою ліфта в чотириповерховому будинку. Станами автоматів є номери поверхів. Вхідний сигнал — номер потрібного поверху, вихідний сигнал — напрямок руху (вгору, вниз, не рухатись).

7.2. Автоматне відображення

Нехай у процесі функціонування заданого автомата A вхідний сигнал у момент автоматного часу $t = 1$ дорівнює $x_{i_1} \in X$, у наступний момент $t = 2$ — $x_{i_2} \in X$ і т. д., а в момент $t = k$ на вхід автомата A подається сигнал $x_{i_k} \in X$. Інакше кажучи, $x(1) = x_{i_1}$, $x(2) = x_{i_2}$, ..., $x(k) = x_{i_k}$. У цьому разі будемо говорити, що на вхід автомата A подано ***вхідне слово*** $x_{i_1}x_{i_2}\ldots x_{i_k} \in X^*$.

Для заданого автомата A його функції переходів δ і виходів λ можна природно поширити з множини $U \times X$ на множину $U \times X^*$, даючи змогу визначати стан і вихідний сигнал в автоматі A після подання на його вхід довільного вхідного слова $p = x_{i_1}x_{i_2}\ldots x_{i_k} \in X^*$. Вважаємо, що

$$\delta(a_l, p) = \delta(\delta(\ldots(\delta(\delta(a_l, x_{i_1}), x_{i_2}), \ldots), x_{i_{k-1}}), x_{i_k}),$$
$$\lambda(a_l, p) = \lambda(\delta(\ldots(\delta(\delta(a_l, x_{i_1}), x_{i_2}), \ldots), x_{i_{k-1}}), x_{i_k}).$$

Іноді ***розширені функції переходів*** і ***виходів*** автомата позначають особливим чином, наприклад δ^* і λ^*, але ми цього не робитимемо й залишимо для розширених функцій ті самі позначення, що й для основних.

Розширену функцію переходів δ можна також означити індуктивно:

1) $\delta(a_l, x)$ для $x \in X$ визначають за таблицею переходів автомата A;

2) для довільного слова $p \in X^*$ та довільного вхідного сигналу $x \in X$ $\delta(a_l, px) = \delta(\delta(a_l, p), x)$.

Спираючись на останнє означення, розширену функцію виходів λ можна означити співвідношенням

$$\lambda(a_l, px) = \lambda(\delta(a_l, p), x), \; p \in X^*, x \in X. \tag{7.2}$$

Для порожнього слова $e \in X^*$ вважають, що $\delta(a_l, e) = a_l$ і $\lambda(a_l, e) = e$.

Нехай A/a_1 — ініціальний автомат. Для довільного вхідного слова $p = x_{i_1}x_{i_2}\ldots x_{i_k}$ означимо відповідне вихідне слово $q \in Y^*$ так:

$$q = \lambda(a_1, x_{i_1})\lambda(a_1, x_{i_1}x_{i_2})\lambda(a_1, x_{i_1}x_{i_2}x_{i_3})\ldots\lambda(a_1, x_{i_1}x_{i_2}\ldots x_{i_k}).$$

Так означена відповідність між вхідними словами p та вихідними словами q називається ***автоматним відображенням***, що індукується (ініціюється, реалізується) автоматом A/a_1; його позначають φ_A. Рівносильним означенням автоматного відображення $\varphi_A\colon X^* \to Y^*$ є такі рекурентні співвідношення:

$$\varphi_A(e) = e, \quad \varphi_A(px) = \varphi_A(p)\lambda(a_1, px), \; p \in X^*, x \in X. \tag{7.3}$$

Автоматне відображення часто називають також ***поведінкою***, або *зовнішньою поведінкою*, автомата.

Автоматне відображення має такі дві важливі властивості, що випливають безпосередньо з його означення:

1) $|p| = |\varphi_A(p)|$ для будь-якого слова $p \in X^*$. Через $|w|$ позначено довжину слова w;

2) якщо $p = p_1p_2$ та $\varphi_A(p) = q_1q_2$, де $|p_1| = |q_1|$, то $q_1 = \varphi_A(p_1)$.

Назвемо сформульовані властивості ***умовами автоматності*** відображення φ_A.

Поняття автоматного відображення можна узагальнити, аналогічно означивши відповідність між вхідними і вихідними словами, індуковану автоматом A/a_i, тобто автоматом A, що починає свою роботу зі стану a_i. Позначатимемо таке автоматне відображення через φ_A^i.

Із наведених означень і умов автоматності випливає така важлива властивість автоматних відображень φ_A^i: якщо $p = p_1p_2 \in X^*$ й $a_{i_1} = \delta(a_i, p_1)$, то

$$\varphi_A^i(p) = \varphi_A^i(p_1p_2) = \varphi_A^i(p_1)\varphi_A^{i_1}(p_2). \tag{7.4}$$

Усі наведені означення можна наочно проілюструвати за допомогою графа автомата. Якщо зафіксувати якийсь стан $a_i \in U$ в автоматі A, то будь-яке слово $p = x_{i_1}x_{i_2}\ldots x_{i_k} \in X^*$ однозначно визначає шлях довжини k, що веде з вершини a_i і складається з k дуг, позначених послідовно вхідними сигналами x_{i_1}, x_{i_2}, ..., x_{i_k}. Тоді стан $\delta(a_i, p)$ — це остання вершина цього шляху, $\lambda(a_i, p)$ — вихідний сигнал, яким позначено останню його дугу, а $\varphi_A^i(p)$ — слово, утворене з послідовності k вихідних сигналів, які написані на k дугах цього шляху.

Приклад 7.4. Для автомата дорожнього руху D з прикладу 7.1 і для вхідних слів $p_1 = x_1x_1x_2$ і $p_2 = x_2x_2x_1x_2$ маємо: $\delta(a_1, p_1) = a_3$, $\delta(a_3, p_1) = a_5$, $\delta(a_4, p_2) = a_1$, $\lambda(a_1, p_1) = y_3$, $\lambda(a_4, p_2) = y_1$, $\lambda(a_2, p_2) = y_3$. Крім того, $\varphi_D(p_1) = y_2y_1y_3$, $\varphi_D(p_2) = y_3y_4y_2y_3$, $\varphi_D^2(p_1) = y_1y_2y_3$, $\varphi_D^4(p_2) = y_2y_3y_4y_1$.

7.2. Задачі і вправи

1. Дати індуктивне означення розширеної функції виходів.

2. За допомогою методу математичної індукції довести умови автоматності для автоматного відображення φ_A.

3. Методом математичної індукції за довжиною слова p_2 довести властивість (7.4) автоматного відображення.

4. Знайти значення автоматного відображення φ_S на даному вхідному слові для автомата S із прикладу 7.2:

(а) $\varphi_S^1(x_1x_2x_1x_2x_2)$;

(б) $\varphi_S^1(x_2x_1x_1x_4x_2x_2)$;

(в) $\varphi_S^2(x_2x_4x_2x_4x_2x_3x_2x_1x_2)$.

5. Знайти значення автоматного відображення φ_Z на даному вхідному слові для автомата Z із прикладу 7.3:

(а) $\varphi_Z^0(x_2x_2x_1x_2x_1)$;

(б) $\varphi_Z^2(x_3x_2x_3x_2x_1x_1)$;

(в) $\varphi_Z^1(x_3x_2x_1x_2x_2x_3x_2x_3x_2)$.

6. Для автомата $A = (X, Y, U, \delta, \lambda)$ та вхідного слова $w = x_{i_1}x_{i_2}\ldots x_{i_k}$ означимо матрицю $M(w)$ так: елемент i-го рядка та j-го стовпчика $m_{ij}(w)$ матриці $M(w)$ дорівнює 1, якщо $a_j = \delta(a_i, w)$, і $m_{ij}(w) = 0$ в іншому разі. Довести, що $M(w) = M(x_{i_1})M(x_{i_2}) \cdot \ldots \cdot M(x_{i_k})$.

7. Довести, що для автомата $A = (X, Y, U, \delta, \lambda)$ та довільного вхідного слова $w = x_{i_1}x_{i_2}\ldots x_{i_k}$ елемент i-го рядка та j-го стовпчика t_{ij} добутку матриць переходів $M(x_{i_1})M(x_{i_2}) \cdot \ldots \cdot M(x_{i_k})$ дорівнює 1, якщо у графі автомата A існує шлях, що відповідає слову w та веде з вершини з номером i у вершину з номером j, і $t_{ij} = 0$, в іншому разі.

7.3. Гомоморфізм, ізоморфізм і невідрізнюваність автоматів

Нехай $A_1 = (X_1, Y_1, U_1, \delta_1, \lambda_1)$ і $A_2 = (X_2, Y_2, U_2, \delta_2, \lambda_2)$ — скінченні автомати, й існують три відображення: $\gamma_1: X_1 \to X_2$, $\gamma_2: Y_1 \to Y_2$, $\gamma_3: U_1 \to U_2$. Сукупність відображень $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ називається (X, Y, U)-***гомоморфізмом*** автомата A_1 в автомат A_2, якщо для будь-яких $x \in X_1$ і $a \in U_1$ виконуються умови

$$\gamma_3(\delta_1(a, x)) = \delta_2(\gamma_3(a), \gamma_1(x)), \quad \gamma_2(\lambda_1(a, x)) = \lambda_2(\gamma_3(a), \gamma_1(x)). \qquad (7.5)$$

Якщо крім того всі три відображення $\gamma_1, \gamma_2, \gamma_3$ сюр'єктивні, то вони задають (X, Y, U)-гомоморфізм автомата A_1 *на* автомат A_2. Автомат A_2 називається (X, Y, U)-гомоморфним образом автомата A_1.

Якщо відображення $\gamma_1, \gamma_2, \gamma_3$ взаємно однозначні й виконуються умови (7.5), то сукупність $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ називають (X, Y, U)-***ізоморфізмом*** автомата A_1 на автомат A_2. Автомати, для яких існує (X, Y, U)-ізоморфізм, називають (X, Y, U)-*ізоморфними*.

Означаючи гомоморфізм та ізоморфізм для ініціальних автоматів, вважають, що відображення γ_3 переводить початковий стан одного автомата в початковий стан іншого.

Уведені поняття мають для автоматів той самий смисл, що й для алгебр або графів. Зокрема, автомати A_1 і A_2 ізоморфні, якщо вхідні й вихідні сигнали та стани автомата A_1 можна бієктивно перейменувати так, що таблиці переходів і виходів автомата A_1 перетворяться

в таблиці переходів і виходів автомата A_2. Аналогічно, граф автомата A_1, ізоморфного автомату A_2, після відповідного перейменування вершин та вхідних і вихідних символів на дугах стає графом автомата A_2.

Якщо одне чи більше відображень з гомоморфізму γ тотожні, то їх виключають із γ. Відповідно змінюється й назва гомоморфізму: якщо тотожним є γ_1, то одержимо (Y, U)-гомоморфізм, якщо тотожним є відображення γ_2, то маємо (X, U)-гомоморфізм і т. д. У теорії автоматів найчастіше зустрічається U-гомоморфізм, коли $X_1 = X_2$, $Y_1 = Y_2$ та відображення γ_1 і γ_2 тотожні. U-гомоморфізм називають просто ***гомоморфізмом*** автоматів. Аналогічні позначення та назви використовують також для різних варіантів ***ізоморфізму*** автоматів.

Отже, у разі гомоморфізму (ізоморфізму) автоматів можна вважати, що $\gamma = \gamma_3$, і тоді умови (7.5) мають такий вигляд: для довільних $x \in X_1$ і $a \in U_1$ виконуються співвідношення

$$\gamma(\delta_1(a, x)) = \delta_2(\gamma(a), x), \quad \lambda_1(a, x) = \lambda_2(\gamma(a), x). \qquad (7.6)$$

Розглянемо пару автоматів $A = (X, Y, U_1, \delta_1, \lambda_1)$ і $B = (X, Y, U_2, \delta_2, \lambda_2)$, які мають однакові вхідні й вихідні алфавіти. Отже, усі автоматні відображення, які індукують автомати A і B, мають однакові типи, тобто області визначення та значень цих відображень збігаються. Це дає змогу ввести для автоматів цього типу такі означення.

Стан $a_i \in U_1$ автомата A та стан $b_j \in U_2$ автомата B називаються ***невідрізнюваними***, якщо для довільного слова $p \in X^*$ виконується рівність $\varphi_A^i(p) = \varphi_B^j(p)$.

Наведене означення можна застосовувати й тоді, коли $A = B$. У цьому разі вводять відношення невідрізнюваності для різних станів того самого автомата A: стани $a_i, a_j \in U_1$ автомата A називають *невідрізнюваними*, якщо $\varphi_A^i(p) = \varphi_A^j(p)$ для будь-якого $p \in X^*$.

Автомати A і B називаються ***невідрізнюваними***, якщо для будь-якого стану $a \in U_1$ автомата A існує невідрізнюваний стан $b \in U_2$ автомата B і, навпаки, для будь-якого стану $d \in U_2$ автомата B існує невідрізнюваний стан $c \in U_1$ автомата A. Ініціальні автомати невідрізнювані, коли невідрізнювані їх початкові стани.

Невідрізнюваність автоматів означає, що будь-яке автоматне відображення (поведінку), що реалізує один з автоматів, можна реалізувати іншим автоматам. Інакше кажучи, невідрізнювані автомати за своєю зовнішньою поведінкою подібні між собою.

Неважко переконатися, що відношення невідрізнюваності H рефлексивне, симетричне та транзитивне, й, отже, є відношенням еквівалентності (на множині станів чи на множині автоматів). У зв'язку з цим часто в літературі з теорії автоматів невідрізнюваність називають еквівалентністю, тобто використовують поняття "***еквівалентні стани***", "***еквівалентні автомати***". Термінологічно це не зовсім зручно й не зовсім коректно, бо назву властивості відношення використовують як назву самого відношення. Однак у теорії автоматів ці терміни стали загальноприйнятими, тому в подальшому будемо говорити про еквівалентні стани й еквівалентні автомати, маючи на увазі відношення невідрізнюваності.

Нижченаведена теорема встановлює зв'язок між гомоморфізмом і невідрізнюваністю автоматів.

Теорема 7.1. Якщо існує гомоморфізм γ автомата $A_1 = (X, Y, U_1, \delta_1, \lambda_1)$ на автомат $A_2 = (X, Y, U_2, \delta_2, \lambda_2)$, то автомати A_1 і A_2 невідрізнювані.

Доведення. Для доведення теореми достатньо показати, що для будь-якого стану $a \in U_1$ стани a та $\gamma(a)$ невідрізнювані між собою. Тоді внаслідок сюр'єктивності відображення γ для кожного стану $a' \in U_2$ існує принаймні один стан $a \in \gamma^{-1}(a')$ автомата A_1 такий, що стани $a' = \gamma(a)$ й a невідрізнювані. Отже, автомати A_1 і A_2 будуть невідрізнювані.

Невідрізнюваність станів a та $\gamma(a)$ доведемо методом математичної індукції за довжиною вхідного слова p.

Позначимо через φ_1 і φ_2 автоматні відображення, які індукують відповідно автомати A_1/a і $A_2/\gamma(a)$.

1. База індукції. Якщо $|p| = 0$, тобто $p = e$, то з означення автоматного відображення (7.3) маємо $\varphi_1(e) = \varphi_2(e)$.

Якщо ж $|p| = 1$, тобто $p = x \in X$, то зі співвідношень (7.3) і (7.6) маємо $\varphi_1(x) = \lambda_1(a, x) = \lambda_2(\gamma(a), x) = \varphi_2(x)$. Крім того, $\gamma(\delta_1(a, x)) = \delta_2(\gamma(a), x)$.

2. Крок індукції. Припустімо, що для всіх вхідних слів p довжини k виконується рівність $\varphi_1(p) = \varphi_2(p)$, а також $\gamma(\delta_1(a, p)) = \delta_2(\gamma(a), p)$. Тоді, враховуючи припущення індукції та співвідношення (7.3), (7.2) і (7.6), для довільного слова $px \in X^*$ ($x \in X$) довжини $k + 1$ маємо:

$$\varphi_1(px) \overset{(7.3)}{=} \varphi_1(p)\lambda_1(a, px) \overset{(7.2)}{=} \varphi_2(p)\lambda_1(\delta_1(a, p), x) \overset{(7.6)}{=} \varphi_2(p)\lambda_2(\gamma(\delta_1(a, p)), x) \overset{(7.6)}{=}$$
$$= \varphi_2(p)\lambda_2(\delta_2(\gamma(a), p), x) \overset{(7.2)}{=} \varphi_2(p)\lambda_2(\gamma(a), px) \overset{(7.3)}{=} \varphi_2(px)$$

(для кращого розуміння перетворень між кожними двома виразами написано номер відповідного співвідношення).

Припущення індукції було використано на другому та четвертому кроках перетворень. Теорему доведено.

Неважко переконатись у справедливості таких тверджень, що випливають із доведеної теореми.

Наслідок 7.1.1. Якщо автомати A_1 і A_2 U-ізоморфні (у нашій термінології — просто ізоморфні), то автомати A_1 і A_2 невідрізнювані.

Наслідок 7.1.2. Якщо для автомата A_1 існують гомоморфізми на автомати A_2 та A_3, то автомати A_2 і A_3 невідрізнювані.

7.3. Задачі і вправи

1. Довести, що відношення невідрізнюваності є відношенням еквівалентності.

2. Довести, що ізоморфні автомати невідрізнювані.

3. Побудувати приклад невідрізнюваних автоматів, які не ізоморфні.

4. Довести наслідок 7.1.2.

7.4. Мінімальний (зведений) автомат

Як і для алгебр, у кожному класі еквівалентних між собою об'єктів означають стандартних представників — так звані *канонічні*, або *нормальні, форми*. Зведенням до канонічної форми можна перевірити (або встановити) еквівалентність або нееквівалентність певних об'єктів даної множини. У теорії автоматів такою канонічною формою є так званий мінімальний автомат.

Для множини (класу) K всіх невідрізнюваних між собою скінченних автоматів ***мінімальним***, або ***зведеним***, автоматом називають такий, що належить цій множині й усі різні стани якого попарно нееквівалентні.

Теорема 7.2. У множині K всіх невідрізнюваних між собою скінченних автоматів існує один і з точністю до ізоморфізму тільки один мінімальний автомат Z. Для кожного автомата $A \in K$ існує гомоморфізм γ на автомат Z. Автомат Z має найменшу можливу кількість станів серед усіх автоматів класу K.

Доведення. Нехай $A = (X, Y, U, \delta, \lambda)$ — довільний автомат з класу K. Розіб'ємо множину станів U автомата A на класи еквівалентності за відношенням невідрізнюваності H. Позначимо одержану фактор-множину U/H через Q.

Якщо стани $a_1, a_2 \in U$ належать одному класу еквівалентності $b \in Q$, то для довільного вхідного сигналу $x \in X$ стани $\delta(a_1, x)$ і $\delta(a_2, x)$ також міститимуться в одному класі еквівалентності $b' \in Q$. Справді, якщо припустити, що $a_i = \delta(a_1, x)$ і $a_j = \delta(a_2, x)$ нееквівалентні, то існує таке слово $p \in X^*$, що $\varphi_A^i(p) \neq \varphi_A^j(p)$. Тоді з $\varphi_A^1(xp) = \lambda(a_1, x)\varphi_A^i(p)$ і $\varphi_A^2(xp) = \lambda(a_2, x)\varphi_A^j(p)$ (див. (7.4)) одержимо $\varphi_A^1(xp) \neq \varphi_A^2(xp)$, що суперечить припущенню про невідрізнюваність станів a_1 і a_2.

Крім того, з означення відношення невідрізнюваності випливає, що для невідрізнюваних станів $a_1, a_2 \in U$ з класу $b \in Q$ та довільного вхідного сигналу $x \in X$ виконується $\lambda(a_1, x) = \lambda(a_2, x) = y' \in Y$.

Усе сказане дає змогу побудувати автомат $Z = (X, Y, Q, \delta', \lambda')$, означивши його функції переходів і виходів такими співвідношеннями: для класу (стану) $b \in Q$ та довільного вхідного сигналу $x \in X$

$$\delta'(b, x) = b' \text{ і } \lambda'(b, x) = y', \tag{7.7}$$

де $b' = [\delta(a, x)]_H$ — клас невідрізнюваності, якому належать стани $\delta(a, x)$, а $y' = \lambda(a, x)$ для всіх $a \in b$. Порівнюючи принципи побудови автомата Z зі схемою побудови фактор-алгебри (див. розд. 2.4), автомат Z доречно було б назвати "фактор-автоматом" автомата A за відношенням невідрізнюваності H.

Позначимо через γ канонічне відображення множини станів U на фактор-множину $Q = U/H$, тобто відображення, яке кожному стану $a \in U$ ставить у відповідність клас еквівалентності $b \in Q$, який містить a. Тоді співвідношення (7.7) можна переписати у вигляді $\delta'(\gamma(a), x) = \gamma(\delta(a, x))$ і $\lambda'(\gamma(a), x) = \lambda(a, x)$, $a \in U$, $x \in X$.

Отже, γ — шуканий гомоморфізм автомата A на автомат Z. За теоремою 7.1 автомати A і Z невідрізнювані, тому $Z \in K$. При цьому кожний стан a автомата A невідрізнюваний від стану $\gamma(a) = b$ такого, що $a \in b$ і b містить усі стани автомата A, невідрізнювані зі станом a.

З останнього зауваження випливає, зокрема, що всі стани автомата Z попарно нееквівалентні. Інакше, якщо припустити, що існує пара різних станів b і b' автомата Z, невідрізнюваних між собою, то, вважаючи, що $b \setminus b' \neq \varnothing$, розглянемо стан $a \in b \setminus b'$. Тоді з того, що стан a невідрізнюваний від стану b і з невідрізнюваності станів b і b' одержимо, що стан a еквівалентний усім станам, які входять у b', однак не міститься в b'. Це суперечить принципу побудови множини Q станів автомата Z. Отже, автомат Z — мінімальний.

Доведемо, що автомат Z має найменшу можливу кількість станів серед усіх автоматів класу K. Припустімо, що у множині K існує автомат W з меншою кількістю станів. Із того, що автомати Z і W належать класу K, випливає, що вони еквівалентні (невідрізнювані) між собою. За означенням невідрізнюваності для кожного стану автомата Z існує еквівалентний йому стан автомата W. Оскільки кількість станів автомата W менша, то остання відповідність неін'єктивна, й існують принаймні два стани b і b' автомата Z, яким відповідає один і той самий еквівалентний їм стан автомата W. Тоді внаслідок транзитивності стани b і b' еквівалентні, що суперечить мінімальності автомата Z.

Нарешті, нехай Z' — інший мінімальний автомат у класі K. Автомати Z і Z' еквівалентні (невідрізнювані), і кількість станів автомата Z' дорівнює кількості станів автомата Z. Справді, якщо припустити, що кількість станів одного з цих автоматів менша, то, повторюючи наведені вище міркування, одержимо суперечність з умовою мінімальності іншого автомата. Невідрізнюваність автоматів Z і Z', а також рівнопотужність множин Q та Q' їх станів дає змогу побудувати взаємно однозначну відповідність η між Q та Q' таку, що пари $(b, b') \in \eta$ складаються з невідрізнюваних між собою станів $b \in Q$ та $b' \in Q'$. Неважко переконатися, що для η виконуються співвідношення (7.6), тому η — ізоморфізм для автоматів Z і Z'. Теорему доведено.

Доведена теорема подає лише інформацію про структуру й основну ідею (принцип) побудови мінімального автомата, однак не містить методу (алгоритму) його побудови. Для того щоб, виходячи з якогось заданого скінченного автомата A, побудувати еквівалентний йому мінімальний автомат Z, потрібно мати конструктивну процедуру, за допомогою якої можна знайти класи еквівалентних (невідрізнюваних) станів автомата A. Ця процедура не випливає безпосередньо з означення відношення невідрізнюваності, бо для підтвердження невідрізнюваності станів a_i і a_j автомата A потрібно встановити рівність відображень φ_A^i і φ_A^j, область визначення X^* яких нескінченна (зліченна).

Наступний розділ присвячено одному з методів побудови мінімального автомата, що має назву методу мінімізації скінченних автоматів.

7.4. Задачі і вправи

1. Довести, що для бієкції η з доведення теореми 7.2 виконуються умови (7.6).

2. Довести, що мінімальні автомати з того самого класу K невідрізнюваних автоматів ізоморфні між собою.

3. Довести справедливість твердження, оберненого до одного з тверджень теореми 7.2: якщо автомат A з класу K всіх невідрізнюваних між собою скінченних автоматів має найменшу можливу кількість станів серед усіх автоматів класу K, то A – мінімальний автомат.

7.5. Алгоритм мінімізації скінченних автоматів

В основі розглядуваного нижче алгоритму мінімізації скінченного автомата лежить теорема 7.2. Сутність алгоритму полягає в послідовній процедурі побудови класів еквівалентних (невідрізнюваних) станів для множини станів заданого автомата $A = (X, Y, U, \delta, \lambda)$.

Назвемо стани a_i, $a_j \in U$ автомата A ***k*-невідрізнюваними**, якщо $\varphi_A^i(p) = \varphi_A^j(p)$ для всіх вхідних слів p довжини k, тобто для всіх $p \in X^k$. Позначимо відношення k-невідрізнюваності через H_k, $k = 1, 2, \ldots$.

Безпосередньо з наведеного означення випливають такі твердження.

Лема 7.1. Якщо стани a_i, $a_j \in U$ k-невідрізнювані, $(a_i, a_j) \in H_k$, то a_i і a_j l-невідрізнювані, $(a_i, a_j) \in H_l$, для $l < k$.

Лема 7.2. Стани a_i, $a_j \in U$ автомата A невідрізнювані тоді й тільки тоді, коли вони k-невідрізнювані для будь-якого $k = 1, 2, \ldots$.

Лема 7.3. Відношення k-невідрізнюваності H_k рефлексивне, симетричне та транзитивне, тобто є відношенням еквівалентності на множині станів U автомата A.

Позначимо фактор-множину U/H_k через Q_k, а класи еквівалентності (класи k-невідрізнюваності) множини Q_k — через $C_i^{(k)}$. Як індекс класу оберемо індекс будь-якого стану $a_i \in U$, що міститься в класі $C_i^{(k)}$. Для однозначності можна брати найменший з усіх таких індексів.

Із леми 7.1 випливає, що

$$C_i^{(k)} \subseteq C_i^{(l)} \quad \text{для } l < k \text{ й усіх } i \leq |U|. \tag{7.8}$$

Доведемо важливу властивість відношення k-невідрізнюваності.

Лема 7.4. l-невідрізнювані стани a_i і a_j автомата A будуть $(l + 1)$-невідрізнюваними тоді й тільки тоді, коли для довільного вхідного сигналу $x \in X$ стани $a_{i_1} = \delta(a_i, x)$ і $a_{j_1} = \delta(a_j, x)$ l-невідрізнювані.

Доведення. Справді, припустімо, що стани a_i і a_j $(l + 1)$-невідрізнювані, однак для якогось $x' \in X$ зазначені стани a_{i_1} і a_{j_1} не є l-невідрізнюваними. З останнього припущення випливає, що існує таке вхідне слово $p \in X^l$, що $\varphi_A^{i_1}(p) \neq \varphi_A^{j_1}(p)$. Тоді зі співвідношення (7.4) маємо $\varphi_A^i(x'p) = \lambda(a_i, x')\varphi_A^{i_1}(p)$ і $\varphi_A^j(x'p) = \lambda(a_j, x')\varphi_A^{j_1}(p)$. Отже, $\varphi_A^i(x'p) \neq \varphi_A^j(x'p)$ для $x'p \in X^{l+1}$, що суперечить припущенню про $(l + 1)$-невідрізнюваність станів a_i і a_j.

Навпаки, нехай для l-невідрізнюваних станів a_i і a_j та довільного $x \in X$ стани $a_{i_1} = \delta(a_i, x)$ і $a_{j_1} = \delta(a_j, x)$ l-невідрізнювані. Тоді за лемою 7.1 вони 1-невідрізнювані, тобто для будь-якого $x \in X$ $\varphi_A^i(x) = \varphi_A^j(x)$, або $\lambda(a_i, x) = \lambda(a_j, x)$. Отже, для довільного вхідного слова $p \in X^l$ і будь-якого $x \in X$ маємо $\varphi_A^i(xp) = \lambda(a_i, x)\varphi_A^{i_1}(p) = \lambda(a_j, x)\varphi_A^{j_1}(p) = \varphi_A^j(xp)$, $xp \in X^{l+1}$, що й доводить $(l + 1)$-невідрізнюваність станів a_i і a_j.

Зі співвідношень (7.8) випливає, що класи k-невідрізнюваності є підкласами (підмножинами) класів l-невідрізнюваності для $l < k$. Інакше кажучи, розбиття множини U на класи $(l + 1)$-невідрізнюваності є підрозбиттям U на класи l-невідрізнюваності. Оскільки множина станів U автомата A скінченна, а значить, і множина її розбиттів скінченна, то обов'язково існує таке число k, що виконуватиметься рівність $C_i^{(k+1)} = C_i^{(k)}$ для всіх i, тобто $Q_k = Q_{k+1}$, чи $U/H_k = U/H_{k+1}$.

Теорема 7.3. Якщо для якогось k $Q_k = Q_{k+1}$ (або $U/H_k = U/H_{k+1}$), то $Q_k = Q$ (або $U/H_k = U/H$), тобто відповідні класи k-невідрізнюваності є шуканими класами невідрізнюваності для множини станів U заданого автомата A.

Доведення. З умови $Q_k = Q_{k+1}$ і леми 7.4 можна дійти висновку, що будь-які стани a_i і a_j з одного класу k-невідрізнюваності є $(k + 1)$-невідрізнюваними, тому для довільного $x \in X$ стани $a_{i_1} = \delta(a_i, x)$ і $a_{j_1} = \delta(a_j, x)$ також належатимуть одному класу k-невідрізнюваності. Водночас за умовою теореми стани a_{i_1} і a_{j_1} належать одному класу $(k + 1)$-невідрізнюваності. Тоді з $(k + 1)$-невідрізнюваності пар станів $(a_i, a_j) \in H_{k+1}$ і $(a_{i_1}, a_{j_1}) \in H_{k+1}$ і леми 7.4 випливає, що стани a_i і a_j $(k + 2)$-невідрізнювані. Отже, $Q_k = Q_{k+2}$.

Повторюючи ці міркування, одержимо, що $Q_k = Q_l$ для всіх $l > k$, тобто з умови $Q_k = Q_{k+1}$ випливає, що будь-які k-невідрізнювані стани a_i й a_j

будуть l-невідрізнюваними для $l > k$. З леми 7.1 одержимо l-невідрізнюваність цих станів для всіх $l < k$. Отже, стани a_i і a_j l-невідрізнювані для будь-якого $l = 1, 2, 3, \ldots$, і за лемою 7.2 вони невідрізнювані (еквівалентні). Теорему доведено.

Підсумовуючи все вищесказане, можна сформулювати такий індуктивний алгоритм знаходження класів еквівалентності на множині станів U заданого скінченного автомата $A = (X, Y, U, \delta, \lambda)$ за відношенням невідрізнюваності H, на якому грунтується загальний алгоритм побудови мінімального автомата Z, еквівалентного автомату A.

Алгоритм знаходження класів еквівалентності на множині U

1. Стани a_i, $a_j \in U$ віднесемо до одного класу 1-невідрізнюваності $C_i^{(1)}$ тоді й тільки тоді, коли $\lambda(a_i, x) = \lambda(a_j, x)$ для будь-якого $x \in X$.

2. Стани a_i і a_j з одного класу l-невідрізнюваності $C_i^{(l)}$ $(l \geq 1)$ віднесемо до одного класу $(l + 1)$-невідрізнюваності $C_i^{(l+1)}$ тоді й тільки тоді, коли для будь-якого $x \in X$ стани $a_{i_1} = \delta(a_i, x)$ і $a_{j_1} = \delta(a_j, x)$ належать тому самому класу l-невідрізнюваності $C_{i_1}^{(l)}$.

У разі ж існування такого сигналу x, що стани $\delta(a_i, x')$ і $\delta(a_j, x')$ належать різним класам l-невідрізнюваності, стани a_i й a_j відносимо до різних класів $(l + 1)$-невідрізнюваності $C_i^{(l+1)}$ і $C_j^{(l+1)}$. При цьому класи $C_i^{(l+1)}$ і $C_j^{(l+1)}$ є підкласами (підмножинами) класу $C_i^{(l)}$. Будемо говорити, що за допомогою вхідного сигналу x' клас l-невідрізнюваності $C_i^{(l)}$ розщеплюється на класи $(l + 1)$-невідрізнюваності $C_i^{(l+1)}$ і $C_j^{(l+1)}$.

3. Якщо для якогось k жоден із класів k-невідрізнюваності не розщеплюється на підкласи, тобто для всіх i $C_i^{(k)} = C_i^{(k+1)}$, або $Q_k = Q_{k+1}$, то алгоритм завершує свою роботу, й одержане розбиття Q_k є шуканим розбиттям Q множини U на класи еквівалентних (невідрізнюваних) станів.

В іншому разі $(Q_k \neq Q_{k+1})$ повторимо крок 2 для розбиття Q_{k+1}. Одержавши множину Q класів еквівалентності станів автомата A, будуємо функцію переходів δ' і функцію виходів λ' мінімального автомата $Z = (X, Y, Q, \delta', \lambda')$ за формулами (7.7). Задачу мінімізації автомата A розв'язано.

Приклад 7.5. Розглянемо процедуру мінімізації автомата A, таблиці переходів і виходів якого подано на рис. 7.6.

За таблицею функції виходів λ одержимо розбиття Q_1 на класи 1-невідрізнюваності, об'єднуючи у класи еквівалентності стани, стовпчики значень для яких у цій таблиці збігаються: $Q_1 = \{C_1^{(1)}, C_2^{(1)}\}$, де $C_1^{(1)} = \{a_1, a_6\}$, $C_2^{(1)} = \{a_2, a_3, a_4, a_5, a_7\}$. Відтак будуємо допоміжну таблицю для розбиття Q_1, замінюючи в таблиці переходів автомата A стани на відповідні класи 1-невідрізнюваності (рис. 7.7).

δ	a_1	a_2	a_3	a_4	a_5	a_6	a_7
x_1	a_2	a_3	a_6	a_3	a_3	a_4	a_3
x_2	a_1	a_5	a_3	a_7	a_4	a_6	a_7
x_3	a_7	a_4	a_3	a_2	a_5	a_7	a_4

λ	a_1	a_2	a_3	a_4	a_5	a_6	a_7
x_1	y_1	y_1	y_1	y_1	y_1	y_1	y_1
x_2	y_2	y_2	y_2	y_2	y_2	y_2	y_2
x_3	y_2	y_1	y_1	y_1	y_1	y_2	y_1

Рис. 7.6

$\delta^{(1)}$	$C_1^{(1)}$		$C_2^{(1)}$				
	a_1	a_6	a_2	a_3	a_4	a_5	a_7
x_1	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_1^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$
x_2	$C_1^{(1)}$	$C_1^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$
x_3	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$	$C_2^{(1)}$

Рис. 7.7

За допомогою одержаної таблиці знайдемо класи 2-невідрізнюваності множини U. Сигнал x_1 розщеплює клас 1-невідрізнюваності $C_2^{(1)}$ на два класи 2-невідрізнюваності $C_2^{(2)} = \{a_2, a_4, a_5, a_7\}$ і $C_3^{(2)} = \{a_3\}$. Клас $C_1^{(2)}$ не розщеплюється, оскільки обидва стовпчики для станів a_1 і a_6 збігаються, тобто $C_1^{(2)} = C_1^{(1)}$. Отже, $Q_2 = \{C_1^{(2)}, C_2^{(2)}, C_3^{(2)}\}$. Аналогічно будуємо таблицю для розбиття Q_2 (рис. 7.8).

Нарешті, при побудові розбиття Q_3 за допомогою останньої таблиці виникає ситуація, коли жоден із класів 2-невідрізнюваності не розщеплюється на підкласи (для всіх станів з одного класу $C_j^{(2)}$ всі стовпчики в таблиці збігаються, $j = 1, 2, 3$). Отже, $Q_3 = Q_2$, тому шукане розбиття Q множини станів U на класи невідрізнюваності має вигляд

$Q = \{\{a_1, a_6\}, \{a_2, a_4, a_5, a_7\}, \{a_3\}\}$. Позначимо одержані класи невідрізнюваності через $b_1 = \{a_1, a_6\}$, $b_2 = \{a_2, a_4, a_5, a_7\}$, $b_3 = \{a_3\}$ та побудуємо таблиці функції переходів δ' і функції виходів λ' шуканого мінімального автомата $Z = (X, Y, Q, \delta', \lambda')$ за допомогою таблиць автомата A та формул (7.7) (рис. 7.9).

$\delta^{(2)}$	$C_1^{(2)}$		$C_2^{(2)}$				$C_3^{(2)}$
	a_1	a_6	a_2	a_4	a_5	a_7	a_3
x_1	$C_2^{(2)}$	$C_2^{(2)}$	$C_3^{(2)}$	$C_3^{(2)}$	$C_3^{(2)}$	$C_3^{(2)}$	$C_1^{(2)}$
x_2	$C_1^{(2)}$	$C_1^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_3^{(2)}$
x_3	$C_2^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_2^{(2)}$	$C_3^{(2)}$

Рис. 7.8

δ'	b_1	b_2	b_3	λ'	b_1	b_2	b_3
x_1	b_2	b_3	b_1	x_1	y_1	y_1	y_1
x_2	b_1	b_2	b_3	x_2	y_2	y_2	y_2
x_3	b_2	b_2	b_3	x_3	y_2	y_1	y_1

Рис. 7.9

Можна розбити множини станів U заданого автомата A на класи невідрізнюваності також за допомогою спеціальної таблиці. Пропонований метод — це інша форма запису основного алгоритму. Опишемо цей метод, ілюструючи його застосування для автомата A з останнього прикладу.

Побудуємо трикутну таблицю, рядки якої відповідають станам $a_2, a_3, \ldots, a_s$, а стовпчики — станам $a_1, a_2, \ldots, a_{s-1}$. Клітинкам таблиці відповідають невпорядковані пари станів $\{a_i, a_j\}$, $i > j$ (рис. 7.10).

Заповнимо ці клітинки за таким правилом. Якщо для станів a_i і a_j існує такий вхідний сигнал $x \in X$, що $\lambda(a_i, x) \neq \lambda(a_j, x)$, то відповідну клітинку перекреслимо хрестиком. Таким чином ми позначаємо пари станів, які не є 1-невідрізнюваними й потрапляють у різні класи 1-невідрізнюваності.

Для автомата A потрібно, наприклад, перекреслити клітинки $\{a_2, a_1\}$, $\{a_3, a_1\}$, $\{a_4, a_1\}$ тощо. Для пар станів $\{a_i, a_j\}$, які є 1-невідрізнюваними, у відповідну клітинку таблиці запишемо всі пари станів $\{a_{i_1}, a_{j_1}\}$, в які переходить автомат A зі станів a_i і a_j під дією кожного $x \in X$: $a_{i_1} = \delta(a_i, x)$, $a_{j_1} = \delta(a_j, x)$ і такі, що $i_1 \neq j_1$, $\{a_i, a_j\} \neq \{a_{i_1}, a_{j_1}\}$.

a_2	×					
a_3	×	$a_6\ a_3$ $a_5\ a_3$ $a_4\ a_3$				
a_4	×	$a_7\ a_5$	$a_6\ a_3$ $a_7\ a_3$ $a_3\ a_2$			
a_5	×	$a_5\ a_4$	$a_6\ a_3$ $a_4\ a_3$	$a_7\ a_4$ $a_5\ a_2$		
a_6	$a_4\ a_2$ $a_7\ a_5$	×	×	×	×	
a_7	×	$a_5\ a_2$	$a_6\ a_3$ $a_3\ a_2$ $a_4\ a_3$	$a_7\ a_2$ $a_4\ a_2$	$a_4\ a_2$ $a_5\ a_4$	
	a_1	a_2	a_3	a_4	a_5	a_6

Рис. 7.10

Одержимо таблицю, що відповідає етапу побудови розбиття Q_1 та частково містить також інформацію, яку раніше можна було одержати з таблиці розбиття Q_1 (рис. 7.7).

Відтак в одержаній таблиці потрібно закреслити клітинки, які містять пари, що відповідають закресленим раніше клітинкам.

У таблиці на рис. 7.10 слід закреслити клітинки для пар станів $\{a_3, a_2\}$, $\{a_4, a_3\}$, $\{a_5, a_3\}$, $\{a_7, a_3\}$, бо вони містять пару $\{a_6, a_3\}$. Після цього таблиця набирає вигляду, поданого на рис. 7.11.

Остання таблиця відповідає етапу побудови розбиття Q_2 на класи 2-невідрізнюваності, а також таблиці розбиття Q_2.

В одержаній таблиці знову закреслимо всі ті клітинки, які містять пари станів, що відповідають закресленим клітинкам. Дістанемо таблицю розбиття Q_3 і т. д. Процес завершиться тоді, коли утвориться таблиця, в якій не можна закреслити жодної клітинки.

У розглянутому прикладі таку властивість має таблиця на рис. 7.11. Усі незакреслені клітинки підсумкової таблиці відповідають парам невідрізнюваних станів.

Для автомата A маємо такі пари невідрізнюваних станів: $\{a_6, a_1\}$, $\{a_4, a_2\}$, $\{a_5, a_2\}$, $\{a_7, a_2\}$, $\{a_5, a_4\}$, $\{a_7, a_4\}$, $\{a_7, a_5\}$. Враховуючи, що відношення невідрізнюваності транзитивне, одержимо класи невідрізнюваності $\{a_1, a_6\}$, $\{a_2, a_4, a_5, a_7\}$. Кожен стан, що не ввійшов у жоден із побудованих класів невідрізнюваності, не має еквівалентних серед станів розглянутого автомата й утворює одноелементний клас невідрізнюваності. Додаючи ці класи, одержимо шукане розбиття Q.

Остаточно для автомата A маємо $Q = \{\{a_1, a_6\}, \{a_2, a_4, a_5, a_7\}, \{a_3\}\}$.

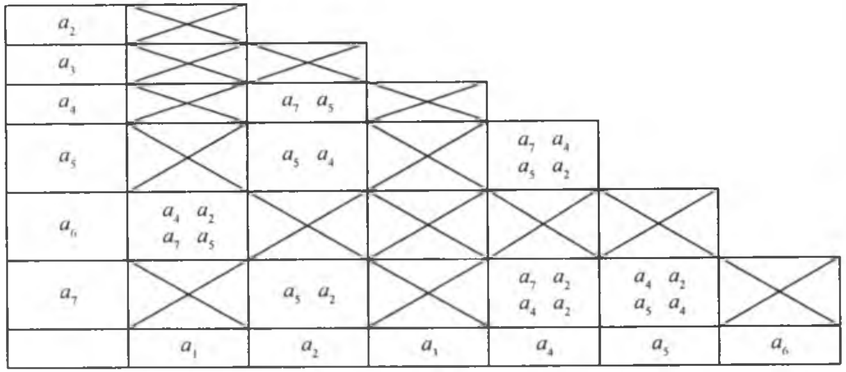


Рис. 7.11

Реалізуючи описану процедуру послідовного закреслення клітинок таблиці, не обов'язково на кожному кроці перерисовувати одержану таблицю, а можна виконати черговий етап безпосередньо для клітинок вихідної таблиці.

7.5. Задачі і вправи

1. Довести леми 7.1–7.3.

2. Довести, що коли в автоматі $A = (X, Y, U, \delta, \lambda)$ для фактор-множин за відношенням k-невідрізнюваності виконується нерівність $Q_k \neq Q_{k+1}$, то $|Q_k| \geq k + 1$.

3. Довести, що стани a_i і a_j автомата $A = (X, Y, U, \delta, \lambda)$ відрізнювані тоді й тільки тоді, коли існує слово $p \in X^*$ довжини не більше $|U| - 1$, для якого $\varphi_A^i(p) \neq \varphi_A^j(p)$.

4. Довести, що для довільного автомата $A = (X, Y, U, \delta, \lambda)$ має місце рівність $Q_{s-1} = Q_s$, де $s = |U|$.

5. Довести, що стани a_i і a_j автомата $A = (X, Y, U, \delta, \lambda)$ невідрізнювані тоді й тільки тоді, коли вони k-невідрізнювані для $k = |U| - 1$.

6. Для довільного натурального n побудувати автомат з n станами, у якому існує пара станів a та b таких, що вони $(n-2)$-невідрізнювані, однак не є $(n-1)$-невідрізнюваними.

7. Довести, що стани a_i і a_j автомата $A = (X, Y, U, \delta, \lambda)$ невідрізнювані тоді й тільки тоді, коли стани $a_{i1} = \delta(a_i, w)$ і $a_{j1} = \delta(a_j, w)$ невідрізнювані для будь-якого $w \in X^*$.

8. Формально обґрунтувати описаний алгоритм знаходження класів невідрізнюваності на множині станів заданого скінченного автомата за допомогою трикутної таблиці.

9. За допомогою обох описаних у розділі методів знайти мінімальний автомат Z для скінченного автомата A, заданого суміщеною таблицею переходів/виходів. Для зручності в таблицях замість стану a_i пишемо тільки його індекс i, а замість вихідного сигналу y_j — індекс j.

(а)

δ/λ	1	2	3	4	5	6	7
x_1	2/1	3/1	6/1	3/1	3/1	4/1	3/1
x_2	1/2	5/2	3/2	7/2	4/2	6/2	2/2
x_3	7/2	4/1	3/1	2/1	5/1	5/2	4/1

(б)

δ/λ	1	2	3	4	5	6	7
x_1	7/2	6/2	7/2	7/2	6/2	6/2	6/1
x_2	1/2	1/1	4/2	3/2	4/1	1/2	4/1
x_3	5/1	3/2	2/1	2/1	1/2	6/1	6/2

10. Перевірити невідрізнюваність двох заданих суміщеними таблицями переходів/виходів скінченних автоматів A_1 і A_2:

(а) A_1:

δ_1/λ_1	1	2	3	4	5	6	7	8	9
x_1	3/1	3/2	3/2	5/2	3/1	1/2	8/2	3/1	3/2
x_2	6/1	7/1	7/2	6/2	7/1	6/2	7/2	4/1	4/1
x_3	3/2	6/2	3/1	2/1	3/2	9/1	2/1	3/2	4/2

A_2:

δ_2/λ_2	1	2	3	4	5	6	7
x_1	7/2	6/2	7/2	7/2	6/2	6/2	6/1
x_2	1/2	1/1	4/2	3/2	4/1	1/2	4/1
x_3	5/1	3/2	2/1	2/1	1/2	6/1	6/2

(б) A_1:

δ_1/λ_1	1	2	3	4	5	6	7	8	9
x_1	3/2	8/1	7/1	2/2	9/2	3/2	3/2	2/2	8/1
x_2	4/1	7/2	7/2	1/1	5/1	1/1	7/1	8/1	8/2
x_3	9/1	1/2	6/2	3/1	9/1	2/1	1/1	4/1	6/2

A_2:

δ_2/λ_2	1	2	3	4	5	6	7
x_1	4/2	7/2	4/2	2/1	4/2	4/2	3/1
x_2	1/1	2/1	6/1	5/2	6/1	3/1	2/2
x_3	7/1	1/1	1/1	1/2	1/1	1/1	1/2

Побудувати відповідні гомоморфні й ізоморфні відображення.

7.6. Дві моделі скінченних автоматів: автомат Мілі та автомат Мура

Скінченний автомат називається ***автоматом Мура***, якщо його функція виходів λ залежить тільки від стану та не залежить від вхідного сигналу, тобто для будь-яких x', $x''\in X$ і $a\in U$ виконується рівність $\lambda(a, x') = \lambda(a, x'')$.

Вилучивши фіктивну змінну, природно вважати функцію виходів автомата Мура унарною функцією типу $U \to Y$. Традиційно цю функцію позначають μ та називають ***функцією відміток***, оскільки вона кожному стану a ставить у відповідність відмітку — вихідний сигнал $y = \mu(a)$.

Розглянута вище загальна модель скінченного автомата називається ***автоматом Мілі***. Обидві моделі названо на честь американських математиків, які одними з перших почали їх систематичні дослідження. Безпосередньо з означення випливає, що автомат Мура — це окремий випадок автомата Мілі.

Способи задання автоматів Мура є модифікацією відповідних способів задання автоматів Мілі.

Так, замість двох таблиць автомата Мілі для задання автомата Мура використовують одну таблицю, яка називається ***відміченою таблицею переходів***. Це звичайна таблиця переходів для функції δ, у якій додатково записано ще один рядок над рядком станів a_1, a_2, …, a_s.

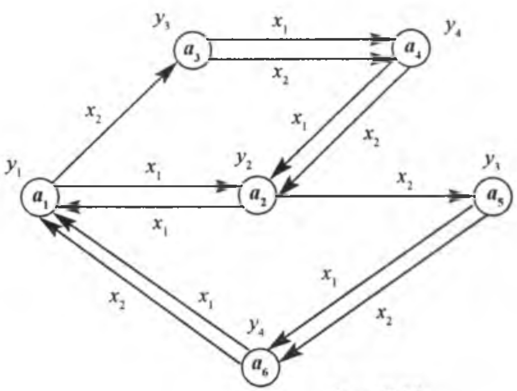
У цьому рядку над кожним зі станів a_i записують вихідний сигнал — значення $\mu(a_i)$ функції відміток для стану a_i.

У графі автомата Мура вихідний сигнал записують не на дугах, а біля вершин. Вершину, що відповідає стану a_i, позначають вихідним сигналом $\mu(a_i)$.

Нарешті, матриці переходів автомата Мура утворюють так само, як і для автомата Мілі, а замість системи матриць виходів використовують одну $s\times n$-матрицю W; елемент w_{ij} i-го рядка та j-го стовпчика матриці W дорівнює 1, якщо $y_j = \mu(a_i)$, і дорівнює 0 — в іншому разі.

Автомат дорожнього руху D з прикладу 7.1 неважко перетворити в автомат Мура. Відмічену таблицю переходів, граф і матрицю виходів автомата D подано на рис. 7.12.

μ	y_1	y_2	y_3	y_4	y_3	y_4
δ	a_1	a_2	a_3	a_4	a_5	a_6
x_1	a_2	a_1	a_4	a_2	a_6	a_1
x_2	a_3	a_5	a_4	a_2	a_6	a_1



$$W = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Рис. 7.12

Поняття розширеної функції переходів залишається для автоматів Мура таким самим, як і для автоматів Мілі. Дещо змінюється друге зі співвідношень (7.3) в означенні автоматного відображення, набираючи вигляду

$$\varphi_A(px) = \varphi_A(p)\,\mu(\delta(a_1, px)). \qquad (7.9)$$

Потребують уточнення також поняття гомоморфізму, ізоморфізму та невідрізнюваності (еквівалентності) станів для автомата Мура.

Так, в означенні (X, Y, U)-гомоморфізму (ізоморфізму) для автоматів Мура $A_1 = (X_1, Y_1, U_1, \delta_1, \mu_1)$ і $A_2 = (X_2, Y_2, U_2, \delta_2, \mu_2)$ другу умову із (7.5) потрібно замінити на умову $\gamma_2(\mu_1(a)) = \mu_2(\gamma_3(a))$, а другу умову із (7.6) — на $\mu_1(a) = \mu_2(\gamma(a))$.

Стан $a_i \in U_1$ автомата Мура $A = (X, Y, U_1, \delta_1, \mu_1)$ і стан $b_j \in U_2$ автомата Мура $B = (X, Y, U_2, \delta_2, \mu_2)$ називаються ***невідрізнюваними***, якщо для довільного вхідного слова $p \in X^*$ виконується рівність $\varphi_A^i(p) = \varphi_B^j(p)$ та, крім того, $\mu_1(a_i) = \mu_2(b_j)$. Після цих уточнень поняття невідрізнюваності (еквівалентності) автоматів Мура та мінімального автомата Мура формулюють так само, як і для автоматів Мілі.

З урахуванням наведених уточнень для автоматів Мура можна сформулювати й довести теореми, аналогічні теоремам 7.1–7.3. Нескладної модифікації потребує також алгоритм мінімізації автоматів. А саме, потрібно ввести поняття 0-невідрізнюваності, вважаючи стани a_i і a_j автомата Мура A 0-невідрізнюваними, якщо $\mu(a_i) = \mu(a_j)$. Відтак на першому кроці мінімізації автомата Мура потрібно розбити множину станів U автомата A на класи 0-невідрізнюваності. Подальші кроки алгоритму залишаються без змін.

Незважаючи на те, що автомат Мура є окремим випадком автомата Мілі, можливості цих двох моделей збігаються.

Теорема 7.4. Для довільного автомата Мілі існує еквівалентний йому автомат Мура.

Доведення. Нехай задано автомат Мілі $A = (X, Y, U, \delta, \lambda)$, де $X = \{x_1, x_2, \ldots, x_m\}$, $U = \{a_1, a_2, \ldots, a_s\}$. Означимо автомат Мура $B = (X, Y, W, \delta', \mu)$ так. Множиною станів автомата B є $W = \{(a_i, x_j) \mid a_i \in U, x_j \in X\} \cup U$, тобто $W = U \times X \cup U$.

Позначимо стан (a_i, x_j) множини W через b_{ij}, а стан $a_i \in U \subseteq W$ — через b_{i0}. Функцію переходів δ' автомата B для стану $b_{ij} \in W$ та вхідного символу $x_l \in X$ означимо так:

1) якщо $j = 0$, тобто $b_{i0} = a_i$, то $\delta'(b_{i0}, x_l) = b_{il} = (a_i, x_l)$;

2) якщо $j \neq 0$, тобто $b_{ij} = (a_i, x_j)$, то $\delta'(b_{ij}, x_l) = b_{kl} = (a_k, x_l)$, де $a_k = \delta(a_i, x_j)$, тобто $\delta'((a_i, x_j), x_l) = (\delta(a_i, x_j), x_l)$.

Функцію відміток μ для станів b_{i0} можна означити довільно, а для решти станів $(j \neq 0)$ $\mu(b_{ij}) = \lambda(a_i, x_j)$.

Доведемо, що стан $a_i \in U$ автомата Мілі A та стан $b_{i0} \in W$ автомата Мура B невідрізнювані, тобто для довільного вхідного слова $p \in X^*$ виконується рівність $\varphi_A^i(p) = \varphi_B^{i0}(p)$.

Доведення проведемо методом математичної індукції за довжиною слова p.

1. База індукції. Якщо $|p| = 0$, тобто $p = e$, то $\varphi_A^i(e) = \varphi_B^{i0}(e) = e$. Якщо ж $|p| = 1$, тобто $p = x_l \in X$, то $\varphi_A^i(x_l) = \lambda(a_i, x_l)$, а $\varphi_B^{i0}(x_l) = \mu(\delta'(b_{i0}, x_l)) = \mu(b_{il}) = \lambda(a_i, x_l)$.

Крім того, для довільного $x_k \in X$

$\delta'(b_{i0}, x_l x_k) = \delta'(\delta'(b_{i0}, x_l), x_k) = \delta'(b_{il}, x_k) = \delta'((a_i, x_l), x_k) = (\delta(a_i, x_l), x_k)$

та $\mu(\delta'(b_{i0}, x_l x_k)) = \lambda(\delta(a_i, x_l), x_k)$.

2. Крок індукції. Припустімо, що для всіх вхідних слів p довжини $n \geq 1$ виконується рівність $\varphi_A^i(p) = \varphi_B^{i0}(p)$, а також що $\mu(\delta'(b_{i0}, px_l)) = \lambda(\delta(a_i, p), x_l)$.

Тоді для довільного слова $px_l \in X^{n+1}$ довжини $n + 1$ маємо $\varphi_A^i(px_l) = \varphi_A^i(p)\lambda(\delta(a_i, p), x_l) = \varphi_B^{i0}(p)\mu(\delta'(b_{i0}, px_l)) = \varphi_B^{i0}(px_l)$.

Теорему доведено.

Тут варто зазначити, що невідрізнюваність автоматів A і B, про яку йдеться в доведеній теоремі, є "однобічною". Для кожного стану автомата Мілі A існує еквівалентний йому стан автомата Мура B, однак обернене твердження не виконується.

Радимо побудувати автомат Мура, еквівалентний автомату Мілі послідовного двійкового суматора з прикладу 7.2.

Отже, при дослідженні можливостей автоматів і реалізації алгоритмів зі скінченною пам'яттю можна користуватися тією чи іншою моделлю скінченного автомата, обираючи ту з них, яка зручніша та більше відповідає розв'язуваній задачі.

7.6. Задачі і вправи

1. Сформулювати детальні означення для автоматного відображення, гомоморфізму й ізоморфізму автоматів і невідрізнюваності (еквівалентності) станів і автоматів для моделі автомата Мура.

2. Сформулювати і довести для автоматів Мура теореми, аналогічні теоремам 7.1–7.3.

3. Сформулювати модифікацію алгоритму мінімізації автомата для моделі автомата Мура.

4. Перевірити невідрізнюваність двох заданих автоматів Мура A_1 і A_2:

(а) A_1

μ_1	1	2	1	1	1	1	1	1	2
δ_1	1	2	3	4	5	6	7	8	9
x_1	3	3	3	1	3	5	8	3	3
x_2	6	4	7	7	4	7	6	6	4
x_3	3	4	3	9	3	2	9	3	7

A_2

μ_2	1	2	1	1	2	1	1
δ_2	1	2	3	4	5	6	7
x_1	7	6	7	7	6	6	6
x_2	4	1	4	4	1	3	3
x_3	2	4	5	2	3	6	6

(б) A_1

μ_1	1	1	1	1	1	1	2	2	1
δ_1	1	2	3	4	5	6	7	8	9
x_1	2	7	8	9	2	3	2	9	7
x_2	5	8	7	4	6	1	8	7	8
x_3	3	6	5	9	3	2	5	1	6

A_2

μ_2	1	2	2	1	2	2	1
δ_2	1	2	3	4	5	6	7
x_1	7	4	4	5	7	4	6
x_2	1	3	5	3	2	6	2
x_3	4	1	1	1	1	1	1

Побудувати відповідні гомоморфні й ізоморфні відображення.

5. Побудувати автомат Мура, еквівалентний заданому автомату Мілі:

(а)

δ	1	2	3	λ	1	2	3
x_1	2	2	1	x_1	1	2	1
x_2	1	3	1	x_2	2	1	1

(б)

δ	1	2	3	λ	1	2	3
x_1	3	2	3	x_1	2	1	1
x_2	1	1	2	x_2	1	1	2

7.7. Основні проблеми теорії автоматів

Щодо зовнішніх умов функціонування розрізняють два типи поведінки та два відповідні типи скінченних автоматів.

Перший тип поведінки — це розглянуте вище і вже знайоме нам перетворення вхідних слів у вихідні, тобто реалізація автоматних відображень. Автомати, що реалізують зазначений тип поведінки, називають ***автоматами-перетворювачами***.

Другий тип поведінки — це розпізнання певних множин вхідних слів. ***Автомат-розпізнавач***, або ***автомат-акцептор***, для будь-якого

вхідного слова p дає змогу визначити, належить це слово певній множині чи ні. У разі позитивної відповіді на поставлене питання кажуть, що слово p *розпізнається*, або *сприймається*, автоматом, а не то вважають, що слово p не розпізнається, або відкидається.

Найпоширенішу модель ініціального автомата-розпізнавача $A = (X, Y, U, \delta, \lambda)$ означають так. Множину станів U розбивають на дві підмножини U_1 і U_2. Вважають, що слово $p \in X^*$ розпізнається автоматом A, якщо $\delta(a_1, p) \in U_1$, в іншому разі ($\delta(a_1, p) \in U_2$) слово p відкидається. Аналогічно, розбивши на дві підмножини множину всіх вихідних сигналів Y, можна означити умову розпізнавання вхідного слова залежно від того, якій із цих підмножин належить вихідний сигнал $\lambda(a_1, p)$ (для автомата Мура — $\mu(\delta(a_1, p))$).

Існує кілька можливостей інтерпретувати зазначені типи автоматів. Наприклад, автомат-перетворювач можна розглядати як пристрій, що перекодовує (перекладає) слова з однієї мови в іншу, або як пристрій, що реалізує алгоритм розв'язання задач певного типу, перетворюючи умови задачі (вхідні слова) у записи відповідних розв'язків (вихідні слова). Можна інтерпретувати автомат-перетворювач і як пристрій, що реалізує певний зворотний зв'язок: автомат одержує інформацію із зовнішнього середовища у вигляді вхідних слів і виробляє відповідні керувальні послідовності у вигляді вихідних слів тощо. Автомат-розпізнавач можна інтерпретувати як аналізатор, що розпізнає (сприймає) певні синтаксичні конструкції (формули, правильні слова, речення, програми тощо).

Кожен із розглянутих типів поведінки автоматів можна звести до іншого. Якщо вхідне слово $p = x_{i1}x_{i2}\ldots x_{ik}$, то послідовність відповідей автомата-розпізнавача на вхідні слова x_{i1}, $x_{i1}x_{i2}$, …, $x_{i1}x_{i2}\ldots x_{ik}$ утворює вихідне слово q — результат автоматного відображення $\varphi_A(p)$. З іншого боку, усі вхідні слова $p \in X^*$ автомата-перетворювача можна класифікувати залежно від останнього символу вихідного слова $\varphi_A(p)$.

Однак попри зазначений тісний зв'язок цих типів поведінки є проблеми, які легко і природно формулюються та розв'язуються для одного типу автоматів й істотно ускладнюються при їх переформулюванні для іншого типу. Тому в теорії автоматів зазначені типи розглядають окремо.

Іноді виділяють окремий випадок першого типу поведінки, реалізований так званими *автономними*, або *породжувальними*, автоматами.

Вхідний алфавіт автономного автомата складається тільки з одного сигналу. Отже, поведінка автомата визначається тільки тими вихідними словами, які здатний породжувати автономний автомат.

Основними проблемами теорії автоматів для кожного з типів поведінки і типів автоматів вважають проблеми аналізу та синтезу.

Проблема аналізу полягає в тому, щоб за заданим автоматом визначити його поведінку та дослідити певні її властивості. Як правило, розв'язання проблеми аналізу принципових труднощів не викликає, оскільки зазвичай разом із заданням автомата наводять правила опису його поведінки.

У ***задачі синтезу*** потрібно, виходячи із заданих умов поведінки, побудувати (синтезувати) автомат, що реалізує цю поведінку. Одна з перших проблем, що виникає в ході розв'язання задачі синтезу, — це дослідження питання, які з умов поведінки взагалі можна реалізувати в автоматах, і зокрема в скінченних автоматах. Відтак постає проблема побудови власне алгоритму синтезу.

7.8. Аналіз і синтез автоматів-перетворювачів

Задачу аналізу, тобто побудову автоматного відображення φ_A ініціального автомата-перетворювача A, можна розв'язати за допомогою формули (7.3) для автоматів Мілі чи формули (7.9) для автоматів Мура.

Нагадаємо дві важливі властивості, які задовольняє автоматне відображення та які названо вище (див. розд. 7.2) *умовами автоматності*:

1) для будь-якого слова $p \in X^*$ $|p| = |\varphi_A(p)|$;

2) якщо $p = p_1p_2$ і $\varphi_A(p) = q_1q_2$, де $|p_1| = |q_1|$, то $q_1 = \varphi_A(p_1)$.

Умови існування розв'язку та метод розв'язання задачі синтезу для автоматів-перетворювачів дає така теорема.

Теорема 7.5. Щоб відображення $\varphi : X^* \to Y^*$ можна було реалізувати деяким автоматом A, необхідно й достатньо, щоб воно задовольняло умови автоматності.

Доведення. Необхідність теореми випливає безпосередньо з означення автоматного відображення, ініційованого автоматом A.

Достатність. Нехай відображення φ вхідних слів у вихідні задовольняє умови автоматності. Розглянемо автомат $A = (X, Y, X^*, \delta, \lambda)$, множина станів якого збігається з множиною всіх слів у алфавіті X.

Початковим станом автомата A вважатимемо порожнє слово e. Функцію переходів δ означимо співвідношенням $\delta(p, x) = px$, $p \in X^*$ і $x \in X$. Значенням функції виходів $\lambda(p, x)$ вважатимемо вихідний сигнал $y_i \in Y$, що є останнім символом слова $\varphi(px)$. Неважко переконатися, що побудований автомат A ініціює відображення φ_A, яке збігається із заданим відображенням φ.

Із доведеної теореми випливає, що умови автоматності рівносильні можливості реалізації відображення φ: $X^* \to Y^*$ в якомусь автоматі. Будемо називати відображення φ типу $X^* \to Y^*$ ***автоматним відображенням***, якщо воно задовольняє умови автоматності.

Побудований у теоремі 7.5 автомат A, що реалізує задане автоматне відображення φ, не є скінченним, оскільки множина X^* його станів нескінченна. Навіть після застосування до нього процедури мінімізації відповідний мінімальний автомат Z може бути, взагалі кажучи, нескінченним. Наступна теорема дає додаткові необхідні й достатні умови, які має задовольняти задане автоматне відображення φ для того, щоб його можна було реалізувати якимось скінченним автоматом.

Нехай задано автоматне відображення φ: $X^* \to Y^*$. Для довільного вхідного слова $p \in X^*$ означимо відображення φ_p: $X^* \to Y^*$, яке називають ***залишковим відображенням*** для відображення φ. Якщо $p' \in X^*$, $p'' = pp'$ і $\varphi(p'') = \varphi(p)q$, то вважатимемо $\varphi_p(p') = q$. Інакше кажучи, результат залишкового відображення φ_p для довільного вхідного слова $p' \in X^*$ можна одержати, утворивши слово $p'' = pp'$, визначивши вихідне слово $\varphi(p'')$ та відкинувши в останньому початковий відрізок довжини $|p|$.

Позначимо через Φ множину всіх залишкових відображень для заданого автоматного відображення φ. Будемо називати *вагою* автоматного відображення φ потужність множини Φ.

Теорема 7.6. Автоматне відображення φ: $X^* \to Y^*$ можна реалізувати деяким скінченним автоматом A тоді й тільки тоді, коли воно має скінченну вагу.

Доведення. Необхідність. Припустімо, що відображення $\varphi = \varphi_A$ ініціює скінченний автомат $A = (X, Y, U, \delta, \lambda)$. Тоді для довільного вхідного слова $p \in X^*$ залишкове відображення φ_p збігається з відображенням φ'_{a_i}, де $a_i = \delta(a_1, p)$ (див. співвідношення (7.4)). Отже, $|\Phi| \leq |U|$; рівність має місце, коли автомат A зведений.

Достатність. Залишкові відображення φ_p, $p \in X^*$ для автоматного відображення φ визначають на множині вхідних слів X^* таке відношення R: $(p', p'')\in R$ для $p', p''\in X^*$ тоді й тільки тоді, коли $\varphi_{p'} = \varphi_{p''}$. Відношення R рефлексивне, симетричне і транзитивне, тому є еквівалентністю. Фактор-множина X^*/R рівнопотужна множині Φ. Якщо автоматне відображення φ має скінченну вагу, то фактор-множина X^*/R скінченна.

Позначимо, як і раніше, через $[p]$ клас вхідних слів з X^*, які еквівалентні слову p. Залишкові відображення з множини Φ, породжені всіма вхідними словами, що еквівалентні слову $p\in X^*$, позначимо $\varphi_{[p]}$, $[p]\in X^*/R$.

Розглянемо скінченний автомат $A = (X, Y, \Phi, \delta, \lambda)$, множиною станів якого є скінченна множина $\Phi = \{\varphi_{[p]} \mid [p]\in X^*/R\}$ залишкових відображень для φ. Початковим станом автомата A є $\varphi_{[e]}$, де $[e]$ — клас слів, еквівалентних порожньому слову $e\in X^*$. Означимо функцію переходів автомата A співвідношенням $\delta(\varphi_{[p]}, x) = \varphi_{[px]}$. Значенням функції виходів $\lambda(\varphi_{[p]}, x)$ вважатимемо $\varphi_{[p]}(x)$, тобто останній символ слова $\varphi(px)$.

Неважко перевірити, що автомат A ініціює задане автоматне відображення φ. Теорему доведено.

На перший погляд може здатися, що умови автоматності досить жорсткі та значно зменшують клас тих відображень типу $X^* \to Y^*$, які можна реалізувати автоматами. Однак існує нескладний спосіб, за допомогою якого довільне відображення φ типу $X^* \to Y^*$ можна перетворити у відповідне автоматне відображення φ_A. Цей спосіб носить назву *стандартного способу вирівнювання довжин слів* [7]. Він грунтується на доповненні вхідного та вихідного алфавітів X і Y спеціальним символом α, відмінним від усіх символів цих алфавітів. Символ α називають порожнім символом, і його появу на вході чи виході автомата інтерпретують як відсутність будь-якого сигналу.

Припустімо, що задане відображення φ слову $p\in X^*$ ставить у відповідність вихідне слово $q\in Y^*$, при цьому $|p| = n$ і $|q| = m$. Позначимо через p_1 слово $p\alpha^m$ і через q_1 — слово $\alpha^n q$, де α^k — це слово $\alpha\alpha\ldots\alpha$ довжини k. Покладемо $\varphi_A(p_1) = q_1$.

Доозначимо відображення φ_A на всіх початкових відрізках слова p_1 такими співвідношеннями:

$$\varphi_A(\eta_k(p_1)) = \eta_k(\varphi_A(p_1)) = \eta_k(q_1), \quad k = 0, 1, 2, \ldots, m + n. \qquad (7.10)$$

Через $\eta_k(w)$ позначено початковий відрізок слова w довжини k, $0 \le k \le |w|$; вважатимемо, що $\eta_0(w) = e$.

Безпосередньо з правил побудови φ_A випливає, що відображення φ_A задовольняє обидві умови автоматності. Покажемо, що при застосуванні розглянутого способу побудови φ_A не виникає неоднозначності, і φ_A — функціональна відповідність.

Неоднозначність може виникнути тоді, коли одне й те саме слово p' в алфавіті $X \cup \{\alpha\}$ є початковим відрізком двох різних слів p_1 і p_2 в цьому самому алфавіті, й при застосуванні доозначення за формулою (7.10) результати для p' можуть виявитися різними.

Розглянемо дві можливі ситуації.

Припустімо, що слово p' містить принаймні один символ α, тоді $p' = p\alpha^l$, $1 \le l \le m$. Це означає, що слово p_2 має вигляд $p\alpha^k$, а з однозначності вихідного відображення φ випливає, що $p_1 = p_2$.

Припустімо тепер, що слово p' не містить жодної літери α. У такому разі $|p'| \le |p|$ і $\varphi_A(p') = \alpha^r$ для всіх слів, початковим відрізком яких є p'.

Значення запропонованого методу полягає в тому, що він дає змогу для довільного відображення типу $X^* \to Y^*$ (словесного відображення) побудувати відповідне автоматне відображення φ_A, тобто відображення, яке можна реалізувати якимось автоматом.

7.7. Задачі і вправи

1. Побудувати (синтезувати) автомат-перетворювач, що функціонує в алфавітах $X = Y = \{0, 1\}$ і реалізує таке автоматне відображення $\varphi_A^{(1)}(x_{i1}\, x_{i2} \ldots \ldots x_{ik}) = y_{j1} y_{j2} \ldots y_{jk}$:

(а) $y(1) = y(2) = 1$, $y(t + 2) = x(t)$;

(б) $y_{jk} = 1$, якщо $x_{i1}x_{i2}\ldots x_{ik}$ містить l одиниць, $y_{jk} = 0$, в іншому разі.

(в) $y_{jk} = 1$, якщо кількість одиниць у слові $x_{i1}\, x_{i2} \ldots x_{ik}$ кратна 3, $y_{jk} = 0$, в іншому разі;

(г) $y(2t - 1) = x(2t - 1)$, $y(2t) = x(2t) \oplus y(2t - 1)$;

(д) $y(1) = 1$, $y(t + 1) = x(t) \oplus x(t + 1)$;

(е) $y(3t - 2) = x(3t - 2) \oplus 1$, $y(3t - 1) = x(3t - 2)$, $y(3t) = x(3t) \wedge y(3t - 1)$, $t = 1, 2, \ldots$.

2. Побудувати автомат-перетворювач, що функціонує в алфавітах $X = Y = \{0, 1\}$ та інвертує кожен четвертий символ вхідної послідовності, тобто $y(4k) = \neg x(4k)$ і $y(t) = x(t)$ для $t \ne 4k$, $k = 1, 2, \ldots$.

3. Нехай $A = (X, Y, U, \delta, \lambda)$ — скінченний автомат, $|U| = s$ і $x\in X$.

(а) довести, що вихідна послідовність z, що відповідає вхідній послідовності $w = xxx\ldots$, періодична;

(б) оцінити довжину передперіоду послідовності z через параметри автомата A;

(в) чому дорівнює максимальна довжина періоду послідовності z?

4. Довести, що для скінченного автомата-перетворювача $A = (X, Y, U, \delta, \lambda)$, на вхід якого подано періодичну послідовність вхідних символів $w = xxx\ldots$, $x \in X$, сума довжин періоду та передперіоду відповідної послідовності вихідних символів не перевищує вагу автоматного відображення φ_A.

5. Довести, що скінченний автомат-перетворювач перетворює періодичні послідовності вхідних символів у періодичні послідовності вихідних символів.

6. Нехай $A = (X, Y, U, \delta, \lambda)$ — скінченний автомат-перетворювач, на вхід якого подано періодичну послідовність вхідних символів $vww\ldots w\ldots$ з періодом w довжини k та передперіодом v довжини d, $w, v \in X^*$. Довести, що період відповідної вихідної послідовності q не перевищує pk, де p — вага автоматного відображення φ_A. Оцінити довжину передперіоду послідовності q.

7.9. Зображення подій в автоматах

Наступні розділи присвячено проблемам аналізу та синтезу для автоматів-розпізнавачів. При дослідженні автоматів-розпізнавачів зручно користуватись автоматами Мура. Можна вважати, що стани автомата-розпізнавача Мура A відмічено двома вихідними сигналами 0 і 1, які розбивають множину станів U автомата A на два класи — U_0 і U_1. Зафіксуємо один із цих класів, позначимо його через F і будемо називати стани з класу F ***заключними***, або ***фінальними*** станами автомата A. Оскільки для автоматів-розпізнавачів нас цікавитиме для кожного вхідного слова лише значення розширеної функції переходів на цьому слові, то доцільно розглянути ще одну модель автомата, яку називають автоматом без виходів.

Автомат без виходів $A = (X, U, \delta, a_1, F)$ означають так. Перші три об'єкти X, U і δ мають той самий смисл, що й раніше: $a_1 \in U$ — початковий стан, F — множина заключних станів автомата A, $F \subseteq U$. У найближчих розділах будемо розглядати без додаткових застережень ініціальні скінченні автомати без виходів.

Множину P слів у вхідному алфавіті X, тобто $P \subseteq X^*$, будемо називати ***подією*** в алфавіті X. Цей термін став традиційним у теорії автоматів, хоча він не містить нічого принципово нового: можна було б користуватися просто терміном "множина слів в алфавіті X". Інший термін для цього поняття — *мова*.

Подія P ***зображувана в автоматі*** $A = (X, U, \delta, a_1, F)$ (або автомат A *зображує* подію P), якщо $\delta(a_1, w) \in F$ тоді й тільки тоді, коли $w \in P$. Кожному автомату $A = (X, U, \delta, a_1, F)$ відповідає зображувана в ньому подія P. Її зручно проілюструвати за допомогою графа автомата A: подія P складається з усіх вхідних слів, відповідні шляхи яких у графі автомата A ведуть з вершини a_1 у вершини з множини F.

Подія P називається ***зображуваною*** (в автоматі), якщо існує автомат A, який її зображує. Інші терміни для цього поняття: множина, або мова, яку *визначає*, *допускає* чи *розпізнає* автомат.

Можлива ситуація, коли $a_1 \in F$. Тоді вважатимемо, що події P, зображуваній в автоматі A, належить порожнє слово e. Його не слід плутати з порожньою подією. Автомат A зображує порожню подію, якщо не існує жодного вхідного слова, яке переводить автомат A з початкового стану a_1 в будь-який із заключних станів. Порожню подію будемо позначати символом порожньої множини $\varnothing$.

Теорема 7.7. Будь-яка скінченна множина слів $P = \{w_1, w_2, \ldots, w_k\}$ в алфавіті X (скінченна подія) зображувана в скінченному автоматі.

Доведення. Нехай $Q = \{\eta_i(w_j) \mid i = 0, 1, \ldots, |w_j|; j = 1, 2, \ldots, k\}$, тобто скінченна множина Q складається з усіх початкових відрізків слів з множини P, зокрема $e \in Q$ і $P \subseteq Q$. Розглянемо автомат без виходів $A = (X, Q \cup \{f\}, \delta, e, P)$. Множина станів автомата A — це множина Q, доповнена деяким станом f ($f \notin Q$); початковим станом автомата A є порожнє слово e, а множиною заключних станів — P. Функцію переходів δ автомата A означимо для $w \in Q$ й $x \in X$ так: $\delta(w, x) = wx$, якщо $wx \in Q$, і $\delta(w, x) = f$, якщо $wx \notin Q$. Крім того, $\delta(f, x) = f$ для всіх $x \in X$.

Оскільки $\delta(e, w_i) = w_i$ для $i = 1, 2, \ldots, k$, то автомат A розпізнає всі слова множини P й тільки ці слова.

Приклад 7.6. На рис. 7.13 подано граф автомата A, що зображує подію $P = \{x_1x_2x_1x_1, x_1x_1x_2, x_1x_1x_1x_2, x_1x_1x_1\}$. Нижче виписано всі слова множини Q та відповідні їм стани автомата A. Заключні стани автомата A позначено подвійними колами.

e — a_1,	x_1x_2 — a_4,	$x_1x_2x_1$ — a_7,
x_1 — a_2,	$x_1x_1x_1$ — a_5,	$x_1x_1x_1x_2$ — a_8,
x_1x_1 — a_3,	$x_1x_1x_2$ — a_6,	$x_1x_2x_1x_1$ — a_9.

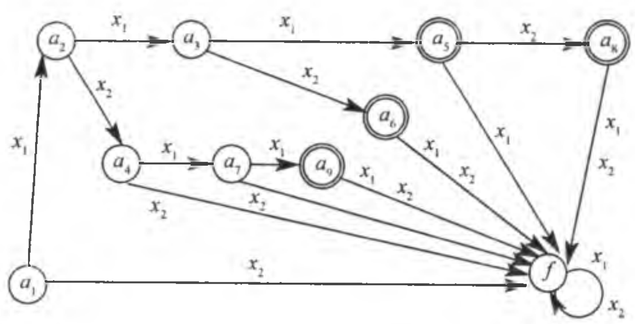


Рис. 7.13

Теорема 7.8. Існують події, не зображувані в скінченних автоматах. Подія P, яка складається з усіх слів w в алфавіті $X = \{x_1, x_2\}$ таких, що слово w містить однакову кількість символів x_1 та x_2, не зображувана в жодному скінченному автоматі.

Доведення. Справедливість першої частини теореми випливає з таких теоретико-множинних міркувань. Зафіксуємо скінченний алфавіт X. Позначимо через M_s сукупність усіх ініціальних автоматів без виходів, множина станів яких містить s елементів: $U = \{a_1, a_2, \ldots, a_s\}$. Множина всіх відображень типу $U \times X \to U$ (тобто множина всіх можливих функцій переходів) скінченна. Кількість можливих варіантів для вибору множини F заключних станів з множини U також скінченна. Отже, множина M_s скінченна.

Множина всіх скінченних ініціальних автоматів без виходів, що функціонують в алфавіті X, дорівнює $M = M_1 \cup M_2 \cup \ldots \cup M_s \cup \ldots$, тому вона є зліченною. Кожен з автоматів множини M зображує одну подію в алфавіті X. Отже, множина подій в алфавіті X, зображуваних у скінченних автоматах, зліченна.

Водночас множина всіх подій в алфавіті X є булеаном зліченної множини X^* та згідно з теоремою 1.7 є незліченною (точніше, множиною потужності континуум). Отже, існують події в алфавіті X, не зображувані в скінченних автоматах. Із теореми 1.6 випливає, що множина таких подій має потужність континуум.

Перший приклад події, не зображуваної в жодному скінченному автоматі, запропонував С.-К. Кліні. Ця подія складається з усіх слів у непорожньому алфавіті X, довжини яких дорівнюють квадратам цілих чисел. У цій теоремі розглянуто інший приклад події, не зображуваної в скінченних автоматах.

Доведення проведемо методом від супротивного. Припустімо, що подія P зображувана в якомусь скінченному автоматі $A = (X, U, \delta, a_1, F)$. Розглянемо послідовність станів $a_{i_k} = \delta(a_1, x_1^k)$, $k = 1, 2, \ldots$ (через w^k позначатимемо слово $ww\!w\ldots w$, що складається з k слів або літер w; вважаємо $w^0 = e$). Зі скінченності множини станів U автомата A випливає, що існують такі числа r і t ($r \neq t$), що $a_{i_r} = a_{i_t}$. Тоді згідно зі зробленим припущенням автомат A розпізнає слово $x_1^r x_2^r$ і відкидає слово $x_1^t x_2^r$, тобто $a' = \delta(a_1, x_1^r x_2^r) \in F$ і $a'' = \delta(a_1, x_1^t x_2^r) \notin F$. Але $\delta(a_1, x_1^r x_2^r) = \delta(\delta(a_1, x_1^r), x_2^r) = \delta(a_{i_r}, x_2^r)$ і $\delta(a_1, x_1^t x_2^r) = \delta(\delta(a_1, x_1^t), x_2^r) = \delta(a_{i_t}, x_2^r)$. Отже, $a' = a''$. Одержана суперечність завершує доведення теореми.

7.8. Задачі і вправи

1. Побудувати автомат-розпізнавач, що зображує подію P:

(а) $P = \{x_1x_1x_1x_2, x_1x_1x_2x_2x_1, x_1x_1x_1x_1x_2, x_1x_1x_1x_2x_1x_1\}$;

(б) $P = \{e, x_1x_2x_1x_1x_2x_2, x_1x_1x_2x_1x_1x_2, x_1x_2x_1x_1\}$.

2. Нехай подія P зображувана в скінченному автоматі A й $e \notin P$. Як побудувати новий скінченний автомат A', що зображує подію $P \cup \{e\}$?

3. Побудувати автомат-розпізнавач, який зображує подію в алфавіті $B = \{0, 1\}$, що складається з усіх слів, котрі починаються та закінчується парами символів 00 або 11, тобто мають вигляд $00w00$ або $11p11$, $w, p \in B^*$.

4. Нехай задано автомат без виходів $A = (X, U, \delta, a_1, F)$, $U = \{a_1, a_2, \ldots, a_s\}$. Позначимо через $M(w)$ матрицю, означену в задачі 7.2(6) для $w \in X^*$, через p — s-вимірний вектор (1, 0, 0, …, 0), а через f — двійковий вектор-стовпчик $(z_1, z_2, \ldots, z_s)^{\mathrm{T}}$, у якому $z_i = 1$, якщо $a_i \in F$, $z_i = 0$ — в іншому разі. Подію називатимемо *матрично зображуваною* в автоматі A, якщо $pM(w)f > 0$ тоді й тільки тоді, коли $w \in P$. Довести, що для того щоб подія P була зображуваною в автоматі A, необхідно й достатньо, щоб вона була матрично зображуваною в A.

5. Будемо говорити, що автомат A розпізнає нескінченну послідовність символів $w = x_{i_1}x_{i_2}\ldots x_{i_k}\ldots$, якщо автомат A зображує подію P, що складається з усіх початкових відрізків послідовності w: $P = \{x_{i_1}, x_{i_1}x_{i_2}, x_{i_1}x_{i_2}x_{i_3}, \ldots\}$. Довести, що неперіодичну послідовність символів w не можна розпізнати жодним скінченним автоматом.

6. Довести, що множина P слів у алфавіті X, довжини яких є точними квадратами, не зображувана в жодному скінченному автоматі.

7. Побудувати скінченний автомат для розпізнавання певної синтаксичної конструкції:

(а) ідентифікатор;

(б) число з фіксованою крапкою;

(в) простий арифметичний вираз;

(г) коментарі.

7.10. Алгебра регулярних подій

У попередньому розділі ми навели приклад класу подій — клас скінченних подій, які завжди зображувані у скінченних автоматах, а також приклад події, незображуваної в скінченних автоматах. Зрозуміло, що в скінченних автоматах зображувані не тільки скінченні події. Бажано було б мати засоби, за допомогою яких можна було б описувати події, зображувані у скінченних автоматах. Один із таких засобів, запропонований С.-К. Кліні й у подальшому розвинутий і вдосконалений В. М. Глушковим, Р. Ф. Мак-Нотоном та іншими, носить назву *мови регулярних виразів*. Ця мова належить до класу так званих алгебричних мов, бо її основою є вирази в певній алгебрі, яка називається *алгеброю регулярних подій*.

Нехай $P_1, P_2 \subseteq X^*$ — події в алфавіті X. Розглянемо три операції над подіями, які називатимемо ***регулярними операціями***.

1. ***Об'єднання*** (іноді *диз'юнкція*) подій $P_1 \cup P_2$ — це звичайне теоретико-множинне об'єднання.

2. ***Множення*** (*конкатенація*) подій P_1P_2 — це подія, яка складається з усіх таких слів w_1w_2, що $w_1 \in P_1$ і $w_2 \in P_2$. Зазвичай операцію множення подій використовують без спеціального позначення. Для зручності посилання на цю операцію позначатимемо її через $\cdot$.

3. ***Ітерацією*** події P називається подія $P^* = \{e\} \cup P \cup P^2 \cup \ldots \cup P^n \cup \ldots = \bigcup_{n=0}^{\infty} P^n$. Тут через P^0 позначено подію $\{e\}$, а через P^n — подію $PP \ldots P$ (n разів). Іноді операцію ітерації події P позначають $\{P\}$.

Звернімо увагу на те, що введене нами вище позначення X^* для множини всіх слів у алфавіті $X = \{x_1, x_2, \ldots, x_m\}$ і позначення X^* операції ітерації для події X мають той самий смисл. Справді, подія $X^* = \left(\{x_1\} \cup \{x_2\} \cup \ldots \cup \{x_m\}\right)^*$ складається з усіх слів (включаючи й порожнє слово) в алфавіті X. Через P^+ позначимо подію $P^* \backslash \{e\}$.

Усі три операції є всюди визначеними на множині подій в алфавіті X. Алгеброю подій в алфавіті X назвемо алгебру $E = \ = \langle \beta(X^*), \{\cup, \cdot, {}^*\} \rangle$, носієм якої є множина всіх подій в алфавіті X (булеан множини X^*), а сигнатура складається з уведених регулярних операцій: двох бінарних (об'єднання та множення) й однієї унарної (ітерації).

Окремо виділимо деякі важливі події. Подію, що містить усі слова в алфавіті X, тобто X^*, назвемо *загальною подією*. Подію, яка не містить жодного слова, назвемо *порожньою*, або *неможливою*, і позначимо символом порожньої множини $\varnothing$.

За допомогою оператора суперпозиції S для операцій алгебри E можна одержувати вирази (*формули* в алгебрі E), які задають певні події в алфавіті X. Щоб зменшити кількість дужок у виразах, домовимося про такі пріоритети для регулярних операцій: спочатку виконується ітерація, відтак — множення і нарешті — об'єднання. Якщо потрібно змінити встановлений порядок виконання операцій, використовують круглі дужки.

Вирази в алгебрі E називають ***рівносильними*** (*еквівалентними*), якщо вони задають ту саму подію. Рівносильність виразів позначають знаком =. Безпосередньо з означень випливають такі властивості регулярних операцій, які ми запишемо у вигляді рівносильних співвідношень для виразів у алгебрі E.

Нехай P, Q і T — довільні події. Тоді

1) $(P \cup Q) \cup T = P \cup (Q \cup T)$,
$(PQ)T = P(QT)$, (*асоціативність*)
2) $P \cup Q = Q \cup P$, (*комутативність об'єднання*)
3) $P(Q \cup T) = PQ \cup PT$,
$(P \cup Q)T = PT \cup QT$, (*дистрибутивність*)
4) $P \cup P = P$, (*ідемпотентність об'єднання*)
5) $(P^*)^* = P^*$, $P^*P^* = P^*$, (*ідемпотентність ітерації*)
6) $PP^* = P^*P$,
7) $P^* = \{e\} \cup PP^* = \left(\{e\} \cup P\right)^*$, (7.11)
$\{e\}P = P\{e\} = P$, $\{e\}^* = \{e\}$, (*властивості події* $\{e\}$)
8) $P\varnothing = \varnothing P = \varnothing$,
$P \cup \varnothing = P$, $\varnothing^* = \{e\}$, (*властивості порожньої події* $\varnothing$)
9) $(P \cup Q)^* = (P^* \cup Q^*)^* = (P^*Q^*)^*$,
10) $(P^* \cup Q)^* = (P \cup Q)^*$,
11) $P^* = \left(\{e\} \cup P \cup P^2 \cup \ldots \cup P^{n-1}\right)(P^n)^*$.

Зазначимо, що операція множення подій некомутативна.

Події $\{x_i\}$, де $x_i \in X$, називають ***елементарними***. До елементарних подій віднесемо також подію $\{e\}$.

Подію P називають ***регулярною***, якщо вона є результатом застосування скінченного числа операцій об'єднання, множення та ітерації до елементарних подій.

Точніше, індуктивне означення регулярних подій таке:

1) елементарні події $\{e\}$ і $\{x_i\}$, $x_i \in X$, $i = 1, 2, \ldots, m$ — регулярні;

2) якщо P_1 і P_2 — регулярні події, то $P_1 \cup P_2$, P_1P_2 та P_1^* — регулярні події;

3) інших регулярних подій, ніж побудовані за правилами 1) і 2), немає.

Позначимо через R клас (множину) всіх регулярних подій в алфавіті X. Клас R замкнений відносно всіх трьох регулярних операцій. Отже, можна розглянути алгебру $E_R = \langle R, \{\cup, \cdot, *\}\rangle$, яка називається ***алгеброю регулярних подій*** і є підалгеброю алгебри подій E. Безпосередньо з означення випливає, що множина елементарних подій є скінченною системою твірних (більше того — базисом) алгебри E_R, тому алгебра регулярних подій — скінченно-породжувана.

З означення випливає також, що кожну регулярну подію можна зобразити (задати) деяким виразом (формулою), що складається зі скінченного числа регулярних подій (операндів) і знаків регулярних операцій. Такі *вирази* називають ***регулярними***. Регулярним виразом для порожньої події вважатимемо символ $\varnothing$.

Регулярні вирази називаються ***рівносильними*** (*еквівалентними*), якщо вони задають ту саму регулярну подію.

Усі наведені вище рівносильні співвідношення (7.11) справедливі в алгебрі регулярних подій E_R, тобто мають місце також для довільних регулярних виразів. Як завжди в алгебрах, ці рівносильності дають змогу виконувати рівносильні перетворення (зокрема, оптимізацію, або спрощення) регулярних виразів. Однак на відміну від вищезгаданих класів еквівалентності (булевих формул, формул в алгебрі Жегалкіна, скінченних автоматів та ін.) для класів еквівалентних регулярних виразів на сьогодні не існує поняття *канонічної форми*, що істотно ускладнює розв'язання проблеми перевірки рівносильності для регулярних виразів.

Будь-яка подія $\{x_{i_1}x_{i_2}\ldots x_{i_k}\}$, що складається з одного слова в алфавіті X^*, регулярна, оскільки вираз $\{x_{i_1}\}\{x_{i_2}\}\ldots\{x_{i_k}\}$, що зображує цю

подію, регулярний. Більше того, будь-яка скінченна подія $P = \{w_1, w_2, \ldots, w_l\}$ є регулярною, і її можна задати виразом $\{w_1\}\cup\{w_2\}\cup \ldots \cup \{w_l\}$. Регулярною є також загальна подія X^*, оскільки $X^* = (\{x_1\}\cup\{x_2\}\cup \ldots \cup\{x_m\})^*$.

Зауважимо, що для зручності та компактності запису регулярних виразів часто випускають фігурні дужки в позначеннях елементарних подій і замість подій (множин) $\{x_i\}$ записують просто x_i, $i = 1, 2, \ldots, m$, а замість $\{e\}$ — e. Іноді для більшої коректності замість $\{x_i\}$ пишуть $\boldsymbol{x}_i$ або $\mathbf{x}_i$.

7.9. Задачі і вправи

1. Нехай P, Q, R і S — події в алфавіті X. Довести, що коли $P \subseteq R$ і $Q \subseteq S$, то $PQ \subseteq RS$. Чи правильним є обернене твердження?

2. Визначити всі можливі події P і Q в алфавіті $\{0, 1\}$, для яких $PQ = \{01, 010, 0101, 0111, 01000, 010111\}$.

3. Довести, що для скінченних подій P і Q виконується нерівність

(а) $|P \cup Q| \le |P| + |Q|$;

(б) $|PQ| \le |P||Q|$.

Для яких P та Q має місце рівність?

4. Виписати всі слова події P^*, довжина яких не перевищує k:

(а) $P = \{ab, b\}$, $k = 5$;

(б) $P = \{a, ab, ba\}$, $k = 7$;

(в) $P = \{aa, aba, baa\}$, $k = 9$.

5. Скільки слів довжини k містить подія, задана регулярним виразом r:

(а) $r = (a \cup b)^*$, $k = 3$;

(б) $r = (aa \cup b)^*$, $3 \le k < 7$;

(в) $r = (a \cup b)^*$, $k = n$;

(г) $r = (aa \cup b)^*$, $k = n$?

6. Навести приклад такої події P, що подія P^* містить:

(а) шестибуквених слів не менше, ніж семибуквених;

(б) шестибуквених слів не менше, ніж восьмибуквених.

Чи існує подія P, для якої P^* містить шестибуквених слів не менше, ніж дванадцятибуквених?

7. Довести, що коли $P \subseteq Q$, то $P^* \subseteq Q^*$.

8. Нехай $P = \{aa, ba\}$ та $Q = \{aa, ba, aaaa\}$. Довести, що $P^* = Q^*$.

9. Нехай $P = \{aa, ba\}$ та $Q = \{aa, ba, aaa\}$. Довести, що $P^* \neq Q^*$ і $P^* \subset Q^*$.

10. Навести приклад таких подій P і Q, що:

(а) $P \subset Q$, однак $P^* = Q^*$;

(б) $P^* = Q^*$, однак $P \not\subset Q$ (символ $\not\subset$ позначає, що не виконується жодне зі співвідношень $\supset$, $\supseteq$, $\subset$, $\subseteq$, $=$).

11. Нехай X — довільний алфавіт. Чи існує нетривіальна подія P в алфавіті X ($P \neq \varnothing$, $P \neq X^*$), для якої виконується таке співвідношення:

(а) $P^n = P$ для всіх $n \in N$;

(б) $P^* = P$?

12. Нехай для непорожньої події P виконується рівність $P^2 = P$. Довести, що:

(а) $e \in P$; (б) $P^* = P$.

13. Навести приклад подій P і Q, для яких події $(P \cup Q)^*$ та $P^* \cup Q^*$ не збігаються. Що можна сказати про події P і Q, якщо виконується рівність $(P \cup Q)^* = P^* \cup Q^*$?

14. Побудувати подію P, для якої P^* містить усі слова в алфавіті $X = \{a, b\}$, довжини яких кратні 3, і тільки такі слова.

15. Довести, що подія P^* замкнена відносно операції конкатенації, тобто для довільних слів $p, q \in P^*$ виконується $pq \in P^*$.

16. Нехай Q — подія, замкнена відносно конкатенації, і $P \subseteq Q$. Довести, що $P^+ \subseteq Q$.

17. Позначимо через $[P]$ замикання події P відносно операції конкатенації (див. розд. 2.5). Довести, що $[P] = P^+$.

18. Довести, що P^+ є найменшою множиною, яка містить P і замкнена відносно конкатенації.

19. Визначити, якою є найменша множина, замкнена відносно конкатенації, що містить події P і Q?

20. Написати регулярний вираз в алфавіті $X = \{0, 1\}$, який задає подію, що складається з усіх таких і тільки таких слів:

(а) у яких кількість символів 0 кратна 3;

(б) які містять рівно 3 символи 1;

(в) які закінчуються на 00 або на 11;

(г) які містять 00 або 11 тільки один раз;

(д) які не містять підслова 111;

(е) які містять парну кількість 1;

(є) у яких сьомим символом є 0;

(ж) у яких третім символом від кінця є 1;

(з) у яких жодні дві 1 чи жодні два 0 не стоять поруч;

(и) у яких нулі утворюють підслова парної довжини.

21. Довести справедливість рівносильних співвідношень (7.11) для виразів у алгебрі E.

22. Описати словами подію, задану таким регулярним виразом:

(а) $a(a \cup b)^*$;

(б) $(aa \cup ab \cup ba \cup bb)^*$;

(в) $\left(aa \cup bb \cup (ab \cup ba)(aa \cup bb)^*(ab \cup ba)\right)^*$;

(г) $(a \cup b)^*a(e \cup bbb)$;

(д) $\left(a(a \cup bb)^*\right)^*$;

(е) $\left(a(aa)^*b(bb)^*\right)^*$;

(є) $\left(b(bb)^*\right)^*\left(a(aa)^*b(bb)^*\right)^*$;

(ж) $\left((a \cup b)a\right)^*$.

23. Побудувати скінченний автомат, що "відшуковує" у вхідному тексті (слові) певний фрагмент, наприклад автомат, що переходить у заключний стан, якщо вхідне слово в алфавіті $X = \{a, b\}$ містить підслово $abba$.

24. Довести такі рівносильні співвідношення для регулярних виразів:

(а) $(a \cup b)^*ab(a \cup b)^* \cup b^*a^* = (a \cup b)^*$;

(б) $(ab)^*a = a(ba)^*$;

(в) $(a^* \cup b)^* = (a \cup b)^*$;

(г) $(a^*b)^*a^* = a^*(ba^*)^*$;

(д) $\left((a \cup bb)^*aa\right)^* = e \cup (a \cup bb)^*aa$;

(е) $(aa)^*(e \cup a) = a^*$;

(є) $(a^*bbb)^*a^* = a^*(bbba^*)^*$;

(ж) $a(ba \cup a)^*b = aa^*b(aa^*b)^*$.

25. Довести чи спростувати твердження про рівносильність таких регулярних виразів:

(а) $(ab \cup a)^*a = a(ba \cup a)^*$;

(б) $(ab \cup a)^*ab = (aa^*b)^*$;

(в) $(a \cup b)^*b = (a^*b)^*$;

(г) $b(ab \cup b)^*a = aa^*b(aa^*b)^*$;

(д) $(a \cup b)^* = a^*(ba^*)^*$;

(е) $a(a \cup b)^* \cup b(a \cup b)^* \cup e = (a \cup b)^*$.

7.11. Події та джерела

Розглянемо спосіб задання подій за допомогою так званих відмічених орієнтованих мультиграфів, або джерел.

Граф називається *відміченим*, якщо його ребрам поставлено у відповідність символи з якогось алфавіту X. Нагадаємо, що граф називається *мультиграфом*, якщо пари його вершин можна з'єднувати більш ніж одним ребром; ці ребра називають *кратними*. Крім того, вважатимемо, що мультиграф може містити петлі.

Як було зазначено вище, у графі ініціального автомата без виходів подія, зображувана цим автоматом, складається з усіх вхідних слів, відповідні шляхи яких у графі автомата ведуть із вершини, що відповідає початковому стану, у вершини, які відповідають заключним станам. Аналогічно, для задання множин слів (подій) можна використовувати відмічені орієнтовані мультиграфи певного виду, які будемо називати джерелами.

Джерелом над алфавітом X називається відмічений орієнтований мультиграф, в якому:

1) виділено початкові та заключні вершини (вершина може бути одночасно і початковою, і заключною);

2) кожну дугу позначено символом з алфавіту X, або порожнім символом e (такі дуги називають ***порожніми***).

Кожне джерело D однозначно визначає певну подію P в алфавіті X, яка складається зі слів у цьому алфавіті, що відповідають усім шляхам, котрі ведуть із початкових вершин у заключні. Якщо шлях містить порожню дугу, то йому відповідає слово $p_1ep_2 = p_1p_2$. Кажуть, що ***джерело D зображує подію*** P.

Джерела називаються ***рівносильними*** (*еквівалентними*), якщо вони зображують ту саму подію.

Граф ініціального автомата без виходів — це окремий випадок джерела.

Для довільного джерела D існує рівносильне йому так зване *двополюсне джерело* D_0, яке має тільки одну початкову й одну заключну вершини. Джерело D_0 можна одержати з D так. Якщо D має декілька початкових вершин, то в D_0 потрібно ввести нову вершину v_0, оголосити її єдиною початковою вершиною джерела D_0 та з'єднати з усіма початковими вершинами джерела D порожніми дугами. Аналогічно, якщо D має декілька заключних вершин, то в D_0 слід увести нову вершину v_F, оголосити її єдиною заключною вершиною джерела D_0 та з усіх заключних вершин джерела D провести порожні дуги у вершину v_F.

Теорема 7.9. Для будь-якої регулярної події P існує двополюсне джерело, яке її зображує.

Доведення. Теорему можна довести індукцією за означенням регулярної події.

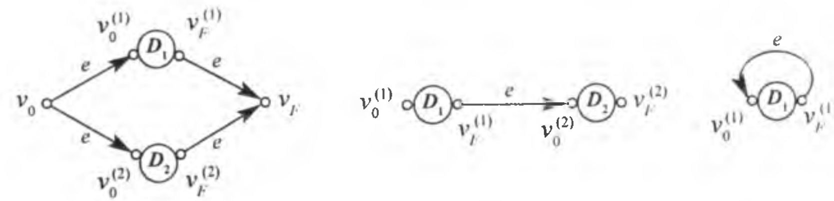
1. *База індукції.* Елементарна подія зображується джерелом, що складається з двох вершин — початкової та заключної, — і дуги між ними, яку відмічено символом цієї події: $x_i \in X$, $i = 1, 2, .., m$, або e.

2. *Крок індукції.* Припустімо, що побудовано двополюсні джерела D_1 і D_2, які зображують відповідно регулярні події P_1 і P_2; початкові вершини D_1 і D_2 — $v_0^{(1)}$ і $v_0^{(2)}$, а заключні — $v_F^{(1)}$ і $v_F^{(2)}$ відповідно. Тоді джерело D з початковою вершиною v_0 і заключною вершиною v_F, яке зображує подію P — результат регулярної операції над подіями P_1 і P_2, можна побудувати як зображено на рис. 7.14.

1) Для $P = P_1 \cup P_2$ джерело D називатимемо "паралельним з'єднанням" D_1 і D_2 (рис. 7.14, *а*). Воно складається з джерел D_1 і D_2 та нових вершин v_0 і v_F: із вершини v_0 потрібно провести порожні дуги в $v_0^{(1)}$ і $v_0^{(2)}$, а з $v_F^{(1)}$ і $v_F^{(2)}$ — порожні дуги у вершину v_F.

2) Для $P = P_1P_2$ джерело D будуємо "послідовним з'єднанням" D_1 і D_2 (рис. 7.14, *б*). Із вершини $v_F^{(1)}$ проводимо порожню дугу у $v_0^{(2)}$. Для джерела D початковою вершиною v_0 оголосимо $v_0^{(1)}$, а заключною вершиною v_F — вершину $v_F^{(2)}$.

3) Якщо $P = P_1^*$, то для побудови джерела D проведемо в D_1 порожню дугу з $v_F^{(1)}$ у $v_0^{(1)}$ (рис. 7.14, *в*). Вершину $v_0^{(1)}$ оголосимо одночасно початковою v_0 та заключною v_F вершиною джерела D.



а *б* *в*

Рис. 7.14

Доведення того, що побудовані таким чином джерела зображують відповідні події, нескладні.

Нехай $P = P_1P_2$ та $w \in P$. Тоді $w = w_1w_2$, де $w_1 \in P_1$ і $w_2 \in P_2$. Оскільки джерело D_1 зображує подію P_1 і D_2 зображує P_2, то в D_1 існує шлях із вершини $v_0^{(1)}$ у $v_F^{(1)}$, що відповідає слову w_1, а в D_2 існує шлях із $v_0^{(2)}$ у $v_F^{(2)}$, що відповідає w_2. Тоді за побудовою в D існує шлях із $v_0 = v_0^{(1)}$ у $v_F = v_F^{(2)}$, що відповідає слову $w = w_1ew_2 = w_1w_2$. Отже, джерело D зображує всі слова події P.

І навпаки, будь-який шлях із вершини v_0 у v_F у джерелі D, якому відповідає слово w, проходить через вершини $v_F^{(1)}$ і $v_0^{(2)}$. Отже, слово w має вигляд w_1w_2, де w_1 — слово, що відповідає шляху з $v_0^{(1)}$ у $v_F^{(1)}$, а w_2 — слово, що відповідає шляху з $v_0^{(2)}$ у $v_F^{(2)}$, тобто $w_1 \in P_1$ і $w_2 \in P_2$. Отже, усі слова, зображувані джерелом D, належать події P.

Доведення для результатів двох інших регулярних операцій — об'єднання й ітерації — аналогічні. Теорему доведено.

В усіх попередніх побудовах було застосовано операцію додавання до джерел нових вершин і нових дуг. Користуючись означенням рівносильності для джерел, можна виконувати в певному розумінні обернені рівносильні перетворення — "спрощення" джерел шляхом вилучення певних дуг і вершин.

Зазначимо деякі очевидні рівносильні перетворення для джерел:

1) вилучення недосяжних вершин разом з усіма дугами, що входять у ці вершини та виходять із них (*недосяжною* вершиною називається вершина джерела, у яку не веде жоден шлях із початкових вершин);

2) вилучення тупикових вершин разом з усіма дугами, що входять у ці вершини і виходять із них (*тупиковою* називають вершину джерела, з якої немає жодного шляху в заключні вершини);

3) вилучення (стягування) деяких порожніх дуг разом з ототожненням ("склеюванням") вершин, які з'єднували ці дуги, тощо.

7.10. Задачі і вправи

1. Виконати повне доведення того, що побудовані за алгоритмом теореми 7.9 джерела зображують відповідні події.

2. Побудувати джерело, що зображує таку подію:

(а) множина таких слів в алфавіті $\{0, 1, 2, \ldots, 9\}$, що остання цифра слова ніде в ньому більше не зустрічається;

(б) множина слів в алфавіті $\{0, 1\}$, у яких рівно два символи 0 і кількість одиниць, що їх розділяють, кратна 4 (зокрема, вона може дорівнювати 0);

(в) множина слів в алфавіті $\{0, 1\}$, в яких серед останніх семи символів є принаймні одна 1.

3. Побудувати джерело, що зображує подію P, задану регулярним виразом:

(а) $ab^* \cup b(ab)^*(a \cup b^*)$;

(б) $(a \cup b^*ab)^*(ba \cup a)^*b$;

(в) $a^*(a \cup b^*)^* \cup ba^*$.

4. Довести, що всі наведені в розділі перетворення для "спрощення" джерел рівносильні. Доповнити цей список правилами для вилучення деяких порожніх дуг.

5. Для всіх джерел з рис. 7.14 виконаємо стягування порожніх дуг (див. розд. 4.10). Довести чи спростувати твердження, що ці перетворення рівносильні.

6. Назвемо *інверсією* (або *транспозицією*) мови L мову L^{-1}, що складається з усіх слів мови L, порядок символів у яких змінено на протилежний. Наприклад, для $L = \{a, aab, abbaa\}$ маємо $L^{-1} = \{a, baa, aabba\}$. Довести, що:

(а) коли існує скінченний автомат, що зображує L, то існує джерело, яке зображує L^{-1};

(б) коли існує джерело, що зображує L, то існує й джерело, яке зображує L^{-1}.

7.12. Синтез автоматів-розпізнавачів

Оскільки граф будь-якого ініціального скінченного автомата без виходів — це окремий випадок джерела, то будь-яка подія, зображувана в автоматі, зображувана також деяким джерелом. Обернене твердження не є очевидним, тому що не кожне джерело є графом скінченного автомата.

Джерело, яке можна інтерпретувати як граф якогось ініціального скінченного автомата без виходів, тобто таке, що має єдину початкову вершину, не містить порожніх дуг і з кожної вершини якого виходить рівно m дуг, позначених символами $x_1, x_2, \ldots, x_m$ алфавіту X, називають ***детермінованим джерелом***.

Наприклад, джерело D, яке зображує задану регулярну подію і побудоване за методом теореми 7.8, навіть після застосування до нього зазначених рівносильних перетворень, узагалі кажучи, не буде детермінованим джерелом.

Наступна теорема пропонує метод перетворення джерела D в рівносильне йому детерміноване джерело.

Теорема 7.10 (про детермінізацію джерела). Для довільного джерела D з n вершинами існує рівносильне йому детерміноване джерело D', яке має не більше ніж 2^n вершин.

Доведення. Множину $W \subseteq V$ вершин джерела D назвемо *замкненою*, якщо з того, що $v \in W$ й існує порожня дуга з v у v', випливає, що $v' \in W$. Для джерела без порожніх дуг будь-яка підмножина множини його вершин замкнена. Позначимо через Q_0 найменшу замкнену множину вершин джерела D, яка містить усі початкові вершини D.

Шукане детерміноване джерело D' будуємо так.

Утворимо всі замкнені підмножини Q_i множини V вершин джерела D (їх кількість не перевищує 2^n — кількості елементів булеана $\beta(V)$) і кожній такій підмножині поставимо у відповідність вершину D'. Початковою вершиною D' оголосимо Q_0, а заключними вершинами D' — усі ті підмножини Q_i, які містять принаймні одну заключну вершину джерела D.

Утворивши вершини джерела D', визначимо його дуги.

Якщо Q_j — найменша замкнена множина, що містить усі вершини, у які з вершин множини Q_i в джерелі D ведуть дуги з відміткою $x_l \in X$, то в джерелі D' проведемо дугу з Q_i в Q_j і позначимо її символом x_l. Якщо ж із всіх вершин множини Q_i в джерелі D не виходить жодної дуги з відміткою x_l, то в D' проведемо дугу з відміткою x_l з вершини Q_i у вершину $\varnothing$, яка відповідає порожній підмножині множини вершин V. Отже, у джерелі D' для кожної вершини Q_i та кожного символу x_l алфавіту X існує рівно одна дуга, що виходить із Q_i, тому D' — детерміноване джерело. Інакше кажучи, D' є графом ініціального скінченного автомата без виходів $A = (X, U, \delta, Q_0, F)$. Описана процедура побудови дуг у D' визначає функцію переходів автомата A: $\delta(Q_i, x_l) = Q_j$.

Доведемо, що джерела D та D' рівносильні й отже, зображують ту саму подію. Позначимо через P і P' події, зображувані відповідно джерелами D і D'.

Для джерел D і D' має місце така властивість: у джерелі D' $\delta(Q_0, w) = Q_j$ для $w \in X^*$ тоді й тільки тоді, коли в D для довільної вершини $v \in Q_j$ існує початкова вершина джерела D $z \in Q_0$, з якої веде шлях у v, що відповідає слову w.

Доведення необхідної та достатньої умов цієї властивості проведемо індукцією за довжиною слова w.

1. Якщо $w = e$, то з умови замкненості маємо $\delta(Q_0, e) = Q_j = Q_0$. Якщо $w = x_l \in X$, то зазначена властивість виконується згідно з правилом побудови дуг у джерелі D' (або згідно з означенням функції переходів δ відповідного автомата A).

2. Припустімо, що властивість виконується для всіх вхідних слів довжини k і $w \in X^k$. Нехай у джерелі D' $\delta(Q_0, wx_l) = Q_j$. Якщо $\delta(Q_0, w) = = Q_i$, то $\delta(Q_i, x_l) = Q_j$.

Згідно з припущенням у джерелі D для довільної вершини $v \in Q_i$ існує шлях із якоїсь початкової вершини $z \in Q_0$ у вершину v, що відповідає слову w. За правилом побудови з того, що у джерелі D' існує дуга з Q_i в Q_j з відміткою x_l, випливає, що в D для довільної вершини $q \in Q_j$ знайдеться вершина $r \in Q_i$, з якої веде дуга в q з відміткою x_l. Отже, у джерелі D є шлях, що відповідає слову wx_l і веде з якоїсь початкової вершини $z \in Q_0$ у вершину $q \in Q_j$.

Аналогічно доведемо обернене твердження.

1. Якщо в D для довільної вершини v з множини Q_j існує дуга з відміткою x_l, що веде з якоїсь початкової вершини $z \in Q_0$ у v, то за побудовою в D' маємо $\delta(Q_0, x_l) = Q_j$.

2. Припустімо, що зазначена властивість виконується для всіх вхідних слів довжини k.

Нехай у джерелі D для всіх вершин $q \in Q_j$ існує початкова вершина $z \in Q_0$, з якої в q веде шлях, що відповідає слову $wx_l \in X^{k+1}$. Розглянемо множину вершин Q_i, у які слово $w \in X^k$ переводить у джерелі D з вершини z. Із припущення індукції маємо в D' рівність $\delta(Q_0, w) = Q_i$. Водночас для вершин множини Q_i існують дуги з відміткою x_l, які ведуть у всі вершини множини Q_j. Це означає, що у джерелі D' виконується рівність $\delta(Q_i, x_l) = Q_j$. Остаточно маємо $\delta(Q_0, wx_l) = Q_j$.

Із доведеної властивості й означення множини F заключних станів джерела D' випливає, що $w \in P'$ тоді й тільки тоді, коли $w \in P$. Отже, $P' = P$. Теорему доведено.

Із теорем 7.9 і 7.10 безпосередньо випливає одна з найважливіших теорем теорії автоматів, яку вперше довів С.-К. Кліні.

Теорема 7.11 (синтезу). Для будь-якої регулярної події P існує скінченний автомат, який її зображує.

Побудова (синтез) шуканого автомата починається з побудови джерела D, яке зображує подію P, методом, запропонованим у теоремі 7.9. Відтак потрібно детермінувати джерело D за допомогою процедури, описаної в доведенні теореми 7.10. На будь-якому з цих кроків можна застосовувати зазначені вище рівносильні перетворення для можливого спрощення джерела. У результаті синтезований автомат може мати менше ніж 2^n станів.

Приклад 7.7. Побудуємо скінченний автомат, який зображує регулярну подію P, задану таким регулярним виразом:

$$(\{x_1\}\{x_2\}^* \cup \{x_3\})(\{x_1\} \cup \{x_2\}\{x_3\})^*\{x_2\}\{x_1\}^*.$$

Спочатку будуємо джерело D, що зображує подію P, за методом теореми 7.9 (рис. 7.15). При цьому деякі “зайві” порожні дуги вилучено. Вершина 1 є початковою, а вершина 8 — заключною.

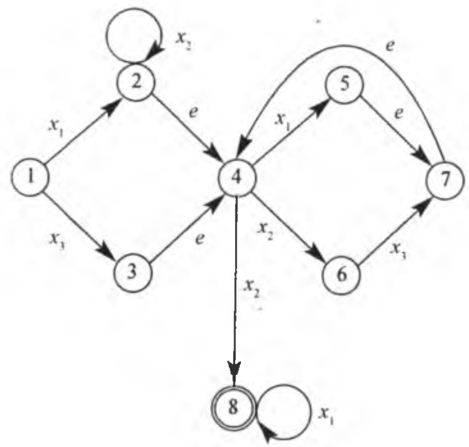


Рис. 7.15

Відтак детермінуємо побудоване джерело згідно з алгоритмом теореми 7.10. При цьому розглядаємо тільки ті підмножини вершин джерела D, які досяжні з початкової вершини $\{1\}$.

Таблицю функції переходів δ детермінованого джерела D', вершинами якого є підмножини вершин джерела D, подано на рис. 7.16, *а*. Оскільки така таблиця у процесі її побудови "росте" внаслідок збільшення кількості станів, то зручно поміняти в ній місцями рядки та стовпчики. Після позначення отриманих підмножин символами з алфавіту $U = \{a_1, a_2, \ldots, a_s\}$ таблиця переходів для D' набирає звичайного вигляду таблиці переходів ініціального скінченного автомата без виходів $A = (X, U, \delta, a_1, F)$ (рис. 7.16, *б*).

Множина заключних станів автомата A складається з тих станів (підмножин), які містять заключну вершину 8 джерела D, тобто $F = \{a_5, a_6, a_7, a_9\}$.

7.11. Задачі і вправи

1. Побудувати скінченний автомат, що зображує подію P, задану регулярним виразом:

(а) $ab^* \cup b(ab)^*(a \cup ba^*)$;

(б) $((a \cup b)^* \cup ab^*)^*(ba \cup a)^*b$;

(в) $a^*(a \cup b^*)^* \cup (ba^* \cup ab)$;

(г) $a^*(a \cup c^*)^* \cup (ba^* \cup ca)b^*$.

δ	x_1	x_2	x_3
$\{1\}$	$\{2, 4\}$	$\varnothing$	$\{3, 4\}$
$\{2, 4\}$	$\{4, 5, 7\}$	$\{2, 4, 6, 8\}$	$\varnothing$
$\{3, 4\}$	$\{4, 5, 7\}$	$\{6, 8\}$	$\varnothing$
$\{4, 5, 7\}$	$\{4, 5, 7\}$	$\{6, 8\}$	$\varnothing$
$\{2, 4, 6, 8\}$	$\{4, 5, 7, 8\}$	$\{2, 4, 6, 8\}$	$\{4, 7\}$
$\{6, 8\}$	$\{8\}$	$\varnothing$	$\{4, 7\}$
$\{4, 5, 7, 8\}$	$\{4, 5, 7, 8\}$	$\{6, 8\}$	$\varnothing$
$\{4, 7\}$	$\{4, 5, 7\}$	$\{6, 8\}$	$\varnothing$
$\{8\}$	$\{8\}$	$\varnothing$	$\varnothing$
$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

а

δ	x_1	x_2	x_3
a_1	a_2	a_{10}	a_3
a_2	a_4	a_5	a_{10}
a_3	a_4	a_6	a_{10}
a_4	a_4	a_6	a_{10}
a_5	a_7	a_5	a_8
a_6	a_9	a_{10}	a_8
a_7	a_7	a_6	a_{10}
a_8	a_4	a_6	a_{10}
a_9	a_9	a_{10}	a_{10}
a_{10}	a_{10}	a_{10}	a_{10}

б

Рис. 7.16

2. Побудувавши відповідні скінченні автомати, перевірити рівність подій P_1 і P_2, заданих регулярними виразами r_1 і r_2:

(а) $r_1 = a^*b \cup (ab)^*a, \quad r_2 = bab^* \cup a^*b$;

(б) $r_1 = (a \cup b)^*a(a \cup b)^*b(a \cup b)^*, \quad r_2 = (a \cup b)^*ab(a \cup b)^*$.

3. Нехай джерело D зображує подію P і $e \notin P$. Як побудувати нове джерело D', що зображує подію $P \cup \{e\}$?

4. Нехай джерело D зображує подію P в алфавіті $\{0, 1\}$ і $01 \notin P$. Як побудувати нове джерело D', що зображує подію $P \cup \{01\}$?

5. Побудувати джерело з $n + 1$ вершиною, що зображує подію P, яка складається з усіх слів у алфавіті $X = \{0, 1\}$, у яких n-й символ від кінця — 1.

6. Довести, що детерміноване джерело, рівносильне джерелу з попереднього прикладу, має щонайменше 2^n станів.

7. Довести, що коли існує скінченний автомат, що зображує подію P, то існує деякий скінченний автомат, який зображує інверсію P^{-1} цієї події.

8. Скільки існує попарно нерівносильних детермінованих джерел із двома вершинами в алфавіті $B = \{0, 1\}$?

7.13. Аналіз автоматів-розпізнавачів

Перейдемо до дослідження проблеми, яку можна вважати протилежною до розглянутої в попередньому розділі, — *проблеми аналізу*, тобто опису подій, зображуваних заданими скінченними автоматами без виходів. Виявляється (і це також уперше довів С.-К. Кліні), що події, зображувані у скінченних автоматах, регулярні. Наступна теорема містить доведення сформульованого твердження й *алгоритм Мак-Нотона–Ямади* побудови відповідного регулярного виразу для події, зображуваної заданим скінченним автоматом.

Теорема 7.12 (аналізу). Будь-яка подія, зображувана в скінченному автоматі, регулярна.

Доведення. Опишемо індуктивну процедуру побудови регулярного виразу для події, зображуваної в заданому скінченному автоматі $A = (X, U, \delta, a_1, F)$, $U = \{a_1, a_2, \ldots, a_s\}$.

Для довільного вхідного слова $w \in X^*$ назвемо *проміжними* стани $\delta(a_i, \eta_l(w))$, $l = 1, 2, \ldots, |w| - 1$.

Позначимо через $P_{ij}^{(k)}$ подію, що складається з усіх вхідних слів w, які переводять автомат A зі стану a_i в стан a_j і таких, що проміжними станами для слова w можуть бути тільки стани з індексами, не більшими від k, тобто тільки стани з множини $\{a_1, a_2, \ldots, a_k\}$.

Доведемо, що $P_{ij}^{(k)}$ — регулярна подія для всіх $i, j = 1, 2, \ldots, s$ і $k = 0, 1, 2, \ldots, s$.

Доведення проведемо індукцією за значенням k.

1. Для $k = 0$ подія $P_{ij}^{(0)}$ складається з усіх слів, які переводять автомат з a_i в a_j і не мають проміжних станів. Такими вхідними словами є або скінченна множина вхідних сигналів $\{x_{l_1}, x_{l_2}, \ldots, x_{l_r}\}$, таких, що $\delta(a_i, x_{l_t}) = a_j$, $t = 1, 2, \ldots, r$, або порожня множина $\varnothing$. Якщо $i = j$, то $e \in P_{ii}^{(0)}$. Отже, $P_{ii}^{(0)}$ — регулярна подія. Регулярний вираз для $P_{ij}^{(0)}$ можна одержати безпосередньо з таблиці переходів автомата A ($i, j = 1, 2, \ldots, s$).

2. Припустімо, що події $P_{ij}^{(k)}$ регулярні для всіх $k = 0, 1, \ldots, t - 1$ і $i, j = 1, 2, \ldots, s$. Доведемо, що

$$P_{ij}^{(t)} = P_{ij}^{(t-1)} \cup P_{it}^{(t-1)}(P_{tt}^{(t-1)})^* P_{tj}^{(t-1)} \qquad (7.12)$$

Нехай $w \in P_{ij}^{(t-1)} \cup P_{it}^{(t-1)}(P_{tt}^{(t-1)})^* P_{tj}^{(t-1)}$. Якщо $w \in P_{ij}^{(t-1)}$, то $w \in P_{ij}^{(t)}$. Якщо ж $w \in P_{it}^{(t-1)}(P_{tt}^{(t-1)})^* P_{tj}^{(t-1)}$, то слово w можна подати у вигляді $w = w_1 w_2 w_3$, де $w_1 \in P_{it}^{(t-1)}$, $w_2 \in (P_{tt}^{(t-1)})^*$ і $w_3 \in P_{tj}^{(t-1)}$. Отже, слово w переводить автомат A зі стану a_i в стан a_j у три етапи: спочатку з a_i в a_t, відтак з a_t в a_t й, нарешті, з a_t в a_j. Для кожного з трьох слів w_1, w_2 і w_3 проміжними є тільки стани з індексами, які не перевищують $t - 1$.

Отже, $w \in P_{ij}^{(t)}$, і ми довели, що подія з правої частини співвідношення (7.12) включається в подію з лівої частини.

Навпаки, нехай $w \in P_{ij}^{(t)}$. Якщо при переведенні автомата A зі стану a_i в стан a_j серед проміжних станів слова w немає стану a_t, то $w \in P_{ij}^{(t-1)}$. Якщо ж при цьому переході використано стан a_t, то процедуру переведення автомата A зі стану a_i в стан a_j можна розбити на три етапи. Спочатку автомат A під дією вхідного слова w_1 переходить зі стану a_i у стан a_t, використовуючи як проміжні лише стани з множини $\{a_1, a_2, \ldots, a_{t-1}\}$. Відтак автомат A під дією слів $w_2^{(1)}, w_2^{(2)}, \ldots, w_2^{(r)}, \ldots$ може кілька разів переходити з a_t в a_t, знову не використовуючи станів з індексами, що перевищують $t - 1$. І нарешті, під дією слова w_3 відбувається перехід з a_t в a_j, при цьому знову з використанням тільки проміжних станів із множини $\{a_1, a_2, \ldots, a_{t-1}\}$. Зрозуміло, що $w_1 \in P_{it}^{(t-1)}$ і $w_3 \in P_{tj}^{(t-1)}$, а всі слова (зокрема, порожнє слово e), які реалізують другий етап переходу, належать множині $\left(P_{tt}^{(t-1)}\right)^*$.
Отже, $w \in P_{it}^{(t-1)}\left(P_{tt}^{(t-1)}\right)^* P_{tj}^{(t-1)}$, і ми довели обернене включення для співвідношення (7.12), а значить, довели й рівність (7.12).

Із припущення індукції про регулярність подій $P_{ij}^{(t-1)}$, $P_{it}^{(t-1)}$, $P_{tt}^{(t-1)}$ і $P_{tj}^{(t-1)}$ та формули (7.12), усі операції якої регулярні, випливає твердження про регулярність події $P_{ij}^{(t)}$.

Якщо $F = \{a_{j_1}, a_{j_2}, \ldots, a_{j_l}\}$, то подію P, зображувану автоматом A, можна записати у вигляді $P = P_{1j_1}^{(s)} \cup P_{1j_2}^{(s)} \cup \ldots \cup P_{1j_l}^{(s)}$ (s — кількість станів автомата A). Із регулярності подій $P_{1j_r}^{(s)}$, $r = 1, 2, \ldots, l$, і регулярності операцій наведеного виразу випливає регулярність події P.

Теорему доведено.

Доведення цієї теореми по суті є алгоритмом аналізу, тобто рекурентною процедурою побудови шуканого регулярного виразу, що задає подію P, зображувану в заданому скінченному автоматі A.

Приклад 7.8. Продемонструємо алгоритм Мак-Нотона–Ямади побудови регулярного виразу для події P, зображуваної в автоматі A, граф якого подано на рис. 7.17. Заключні стани автомата A – a_2 і a_3, $s = 3$.

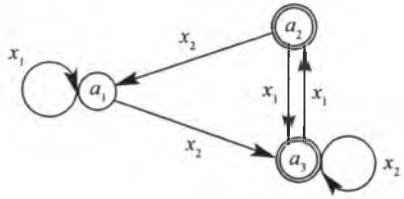


Рис. 7.17

Обчислимо подію $P_{12}^{(3)} \cup P_{13}^{(3)}$. Маємо

$$P_{11}^{(0)} = e \cup x_1, \quad P_{12}^{(0)} = \varnothing, \quad P_{13}^{(0)} = x_2,$$
$$P_{21}^{(0)} = x_2, \quad P_{22}^{(0)} = e, \quad P_{23}^{(0)} = x_1,$$
$$P_{31}^{(0)} = \varnothing, \quad P_{32}^{(0)} = x_1, \quad P_{33}^{(0)} = e \cup x_2.$$

Одержавши вирази для всіх подій $P_{ij}^{(0)}$, $i, j = 1, 2, 3$, побудуємо вирази для $P_{ij}^{(1)}$, $i, j = 1, 2, 3$, за допомогою формули (7.12). При цьому спрощуємо деякі з цих виразів, користуючись тотожностями (7.11):

$P_{11}^{(1)} = e \cup x_1 \cup (e \cup x_1)(e \cup x_1)^*(e \cup x_1) = (x_1)^*$,

$P_{12}^{(1)} = \varnothing \cup (e \cup x_1)(e \cup x_1)^*\varnothing = \varnothing$,

$P_{13}^{(1)} = x_2 \cup (e \cup x_1)(e \cup x_1)^*x_2 = (e \cup ((e \cup x_1)(e \cup x_1)^*)x_2 = (e \cup x_1)^*x_2 =$
$= (x_1)^*x_2$,

$P_{21}^{(1)} = x_2 \cup x_2(e \cup x_1)^*(e \cup x_1) = x_2(x_1)^*$,

$P_{22}^{(1)} = e \cup x_2(e \cup x_1)^*\varnothing = e$,

$P_{23}^{(1)} = x_1 \cup x_2(e \cup x_1)^*x_2 = x_1 \cup x_2(x_1)^*x_2$,

$P_{31}^{(1)} = \varnothing \cup \varnothing(e \cup x_1)^*(e \cup x_1) = \varnothing$,

$P_{32}^{(1)} = x_1 \cup \varnothing(e \cup x_1)^*\varnothing = x_1$,

$P_{33}^{(1)} = e \cup x_2 \cup \varnothing(e \cup x_1)^*x_2 = e \cup x_2$.

Далі

$P_{11}^{(2)} = (x_1)^* \cup \varnothing(e)^*x_2(x_1)^* = (x_1)^*$,

$P_{12}^{(2)} = \varnothing \cup \varnothing(e)^*e = \varnothing$,

$P_{13}^{(2)} = (x_1)^*x_2 \cup \varnothing(e)^*(x_1 \cup x_2(x_1)^*x_2) = (x_1)^*x_2$,

$P_{21}^{(2)} = x_2(x_1)^* \cup e(e)^*x_2(x_1)^* = x_2(x_1)^*$,

$P_{22}^{(2)} = e \cup e(e)^*e = e$,

$P_{23}^{(2)} = x_1 \cup x_2(x_1)^*x_2 \cup e(e)^*(x_1 \cup x_2(x_1)^*x_2) = x_1 \cup x_2(x_1)^*x_2$,

$P_{31}^{(2)} = \varnothing \cup x_1(e)^*x_2(x_1)^* = x_1x_2(x_1)^*$,

$P_{32}^{(2)} = x_1 \cup x_1(e)^*e = x_1$,

$P_{33}^{(2)} = e \cup x_2 \cup x_1(e)^*(x_1 \cup x_2(x_1)^*x_2) = e \cup x_2 \cup x_1x_1 \cup x_1x_2(x_1)^*x_2$.

Нарешті, маємо

$P_{12}^{(3)} = \varnothing \cup (x_1)^*x_2(e \cup x_2 \cup x_1x_1 \cup x_1x_2(x_1)^*x_2)x_1 = (x_1)^*x_2(x_2 \cup x_1x_1 \cup$
$\cup x_1x_2(x_1)^*x_2)^*x_1$,

$P_{13}^{(3)} = (x_1)^*x_2 \cup (x_1)^*x_2(e \cup x_2 \cup x_1x_1 \cup x_1x_2(x_1)^*x_2)^*(e \cup x_2 \cup x_1x_1 \cup$
$\cup x_1x_2(x_1)^*x_2) = (x_1)^*x_2(x_2 \cup x_1x_1 \cup x_1x_2(x_1)^*x_2)^*$.

Отже, подію P, зображувану автоматом A, можна задати регулярним виразом $(x_1)^*x_2(x_2 \cup x_1x_1 \cup x_1x_2(x_1)^*x_2)^*(e \cup x_1)$.

Наведений та інші алгоритми аналізу скінченних автоматів дають, узагалі кажучи, досить складні та громіздкі регулярні вирази. Ці вирази можна певною мірою спростити за допомогою рівносильних перетворень (7.11).

Зауважимо, що в алгебрі регулярних подій існує алгоритм для перевірки рівносильності будь-яких двох регулярних виразів r_1 і r_2. Спочатку, користуючись алгоритмом синтезу, потрібно побудувати автомати A_1 і A_2, які зображують події, задані виразами r_1 і r_2. Відтак, мініімізувавши автомати A_1 і A_2 й порівнявши відповідні мінімальні автомати (канонічні форми автоматів A_1 і A_2), можна підтвердити чи відкинути припущення про рівносильність заданих регулярних виразів.

Однак в алгебрі регулярних подій досі немає нетривіальних методів пошуку *мінімального* (оптимального) регулярного виразу, рівносильного даному, істотно відмінних від тривіального методу повного перебору, який грунтується на сформульованому алгоритмі перевірки рівносильності регулярних виразів. Отже, метод мінімізації регулярних виразів у принципі існує, однак він настільки складний і громіздкий, що недоцільно вести мову про його практичне застосування.

Є різні критерії оцінки оптимальності регулярних виразів, основні з яких — *циклічна глибина* регулярного виразу та його *довжина*. Під циклічною глибиною розуміють максимальну кількість вкладених одна в одну пар ітераційних дужок $(\ )^*$, використовуваних для позначення операції ітерації. Довжиною регулярного виразу вважають кількість його операндів. Розробляють евристичні алгоритми аналізу, які дають змогу для переважної більшості скінченних автоматів одержувати відповідні регулярні вирази, що мало відрізняються від оптимальних.

Як наслідок двох останніх теорем 7.11 і 7.12 справедлива така центральна теорема теорії автоматів.

Теорема 7.13 (Кліні). Подія P зображувана в скінченному автоматі тоді й тільки тоді, коли вона регулярна.

Теорема Кліні визначає ту центральну роль, яку відіграють регулярні події та відповідні регулярні вирази в теорії автоматів. Глибше й детальніше досліджують властивості алгебри регулярних подій $E_R = \langle R, \{\cup, \cdot, {}^*\}\rangle$, розробляють різні критерії, що дають змогу визначати, чи належить певна подія до класу регулярних подій тощо.

Зокрема, виявляється, що множина регулярних подій R замкнена також відносно теоретико-множинних операцій доповнення і перетину. Справді, якщо подію P зображує скінченний автомат $A = (X, U, \delta, a_1, F)$, то подію $\overline{P} = X^* \setminus P$ зображує скінченний автомат $A' = (X, U, \delta, a_1, U \setminus F)$, у якому множину заключних станів утворюють ті й тільки ті стани множини U, які не є заключними в автоматі A.

Відтак, якщо події P_1 і P_2 регулярні, то зі співвідношення $P_1 \cap P_2 = \overline{\overline{P_1} \cup \overline{P_2}}$ та замкненості множини регулярних подій відносно операцій доповнення й об'єднання випливає регулярність події $P_1 \cap P_2$. Наведене співвідношення задає також метод, за допомогою якого, виходячи з джерел D_1 і D_2, що зображують події P_1 і P_2, можна побудувати шукане джерело D, яке зображує регулярну подію $P_1 \cap P_2$. Детермінізувавши джерело D, можна одержати відповідний скінченний автомат.

Наведені факти дають змогу розширити сигнатуру алгебри E_R, додавши до неї операції доповнення та перетину. Існують й інші операції, щодо яких множина регулярних подій R замкнена (наприклад, різниця, симетрична різниця та ін.).

Із замкненості множини регулярних подій R відносно теоретико-множинних операцій об'єднання $\cup$, перетину $\cap$ та доповнення $\overline{}$ випливає, зокрема, що алгебра $\langle R, \{\cup, \cap, \overline{}, \varnothing, X^*\}\rangle$ регулярних подій в алфавіті X є булевою алгеброю.

Цікавою, непростою й дуже важливою є проблема розв'язування рівнянь і систем рівнянь в алгебрі регулярних подій. Наприклад, розглянемо "лінійне" рівняння $X = XP \cup Q$, у якому P і Q — відомі події, а X — шукана подія. Розв'язком цього рівняння є QP^*. Якщо $e \in P$, то остання

подія не єдиний розв'язок рівняння. Неважко перевірити, що в цьому випадку подія $X = (Q \cup R)P^*$ також задовольняє дане рівняння для довільної події R.

Зауважимо, що на методах розв'язання систем рівнянь в алгебрі регулярних подій базується ще один алгоритм аналізу скінченних автоматів, відмінний від алгоритму Мак-Нотона–Ямади.

Нарешті, розглянемо комбінований варіант ініціального скінченного автомата — *розпізнавач-перетворювач* $A = (X, Y, U, \delta, \lambda, a_1, F)$. Це автомат-розпізнавач, у якому означено функцію переходів $\lambda\colon U \times X \to Y$. Отже, автомат A ініціює автоматне відображення $\varphi_A\colon X^* \to Y^*$. Нехай P — подія в алфавіті X, яку зображує автомат A. Назвемо множину слів $L = \{\varphi_A(w) \mid w \in P\}$ подією в алфавіті Y, *породжуваною* автоматом A.

Неважко переконатися, що будь-яка подія L, породжувана ініціальним скінченним автоматом розпізнавачем-перетворювачем A, є регулярною подією в алфавіті Y. Справді, якщо з дуг графа автомата A вилучити всі вхідні відмітки $x_i \in X$, $i = 1, 2, \ldots, m$, то утвориться джерело D, яке зображує подію L в алфавіті Y. Детермінізуючи джерело D, одержимо скінченний автомат, що зображує подію L. Звідси за теоремою аналізу 7.12 випливає, що подія L — регулярна.

Отже, скінченні автомати-перетворювачі перетворюють регулярні події в алфавіті X у регулярні події в алфавіті Y.

7.12. Задачі і вправи

1. За допомогою алгоритму Мак-Нотона-Ямади побудувати регулярний вираз для події, зображуваної в автоматі:

(а)

δ	1	2	3
x_1	2	1	3
x_2	3	2	2

(б)

δ	1	2	3
x_1	1	3	3
x_2	2	1	1

Множина заключних станів: 1) $F = \{1, 3\}$; 2) $F = \{2, 3\}$.

2. Побудувати такі два скінченні автомати, що кожне слово в алфавіті $B = \{0, 1\}$ розпізнається в одному з цих автоматів, однак жодне слово в алфавіті B не розпізнається одночасно обома автоматами.

3. Побудувати (синтезувати) скінченний автомат, що зображує подію P. Подію P задано регулярним виразом:

(а) $ba^* \cup a(ba)^* \cup (aba \cup b^*a^*)$; (б) $a^*(b \cup a^*)^* \cup b^*a$.

4. Позначимо через P і Q події, задані регулярними виразами r і q. Написати регулярний вираз, що задає подію $P \cap Q$:

(а) $r = (a \cup b)^*a$, $q = b(a \cup b)^*$;

(б) $r = (a \cup b)^*a$, $q = (a \cup b)^*aa(a \cup b)^*$;

(в) $r = (a \cup b)b(a \cup b)^*$, $q = b(a \cup b)^*$;

(г) $r = (b \cup ab)^*(a \cup e)$, $q = (a \cup b)^*aa(a \cup b)^*$;

(д) $r = (b \cup ab)^*(a \cup e)$, $q = (a \cup ba)^*a$;

(е) $r = (ab^*)^*a$, $q = b(a \cup b)^*$;

(є) $r = (ab^*)^*a$, $q = a(a \cup b)^*$;

(ж) $r = (aa \cup ab \cup ba \cup bb)^*$, $q = b(a \cup b)^*$.

5. Написати регулярний вираз, що задає подію $P \cap Q$ в алфавіті $B = \{0, 1\}$, де P — множина всіх слів непарної довжини, а Q — це множина всіх слів, які:

(а) починаються з 0;

(б) містять парну кількість нулів;

(в) містять непарну кількість нулів.

6. Побудувати (синтезувати) скінченний автомат, що зображує подію $P_1 \cap P_2$, де події P_1 і P_2 задано регулярними виразами r_1 і r_2: $r_1 = (a \cup b)b(a \cup b)^*$, $r_2 = b(a \cup b)^*$.

7. Довести, що регулярна подія P, зображувана в скінченному автоматі з n станами, нескінченна тоді й тільки тоді, коли вона містить таке слово w, що $n \le |w| < 2n$.

8. Сформулювати ефективну процедуру, що дає змогу для заданого скінченного автомата A перевірити, скінченною чи нескінченною є зображувана в ньому подія.

9. Довести, що для нескінченної регулярної події P будь-яке достатньо довге слово $w \in P$ можна записати у вигляді $w = p_1pp_2$ $(p \ne e)$ так, що для всіх $k = 0, 1, 2, \ldots$ виконуватиметься $p_1p^kp_2 \in P$ (*“лема про накачку”*).

10. Користуючись “лемою про накачку” чи іншими методами, довести, що подія P не регулярна:

(а) $P = \{a^kb^k \mid k = 0, 1, 2, \ldots\}$;

(б) $P = \{ww \mid w \in \{a, b\}^*\}$;

(в) $P = \{a^kb^{k+1} \mid k = 0, 1, 2, \ldots\}$;

(г) $P = \{a^kb^ka^k \mid k = 0, 1, 2, \ldots\}$;

(д) $P = \{a^k \mid k$ — просте число$\}$;

(е) $P = \{a^kb^k \mid k = n^2, n = 0, 1, 2, \ldots\}$;

(є) $P = \{a^{k!} \mid k = 0, 1, 2, \ldots\}$;

(ж) $P = \{a^kb^l \mid k > l\}$.

11. Нехай $P = \{a^k \mid k \ge 0, k$ — не просте число$\}$. Довести, що:

(а) подія P нерегулярна;

(б) подія P задовольняє умову “леми про накачку”.

12. Нехай P — подія в алфавіті X. Означимо відношення R на множині X^*: $(u, v) \in R$ тоді й тільки тоді, коли для будь-якого слова $w \in X^*$ виконується або $uw \in P$ і $vw \in P$, або $uw \notin P$ і $vw \notin P$. Довести, що відношення R — еквівалентність. Довести також, що подія P регулярна тоді й тільки тоді, коли фактор-множина X^*/R скінченна.

13. Довести, що множина регулярних подій в алфавіті X замкнена відносно операції:

(а) різниця; (б) симетрична різниця.

14. Нехай X_1 і X_2 — скінченні алфавіти, $X = X_1 \times X_2$ і $w \in X^*$. Якщо $w = (a_1, b_1)(a_2, b_2)\ldots(a_n, b_n)$, то слова $a_1a_2\ldots a_n$ і $b_1b_2\ldots b_n$ називатимемо відповідно першою та другою проекціями слова w; позначатимемо їх $\mathrm{pr}_1(w)$ та $\mathrm{pr}_2(w)$. Якщо $P \subseteq X^*$, то $\mathrm{pr}_i(P) = \{\mathrm{pr}_i(w) \mid w \in P\}$, $i = 1, 2$. Довести, що для довільної регулярної події R в алфавіті X множина $\mathrm{pr}_i(R)$ є регулярною подією в алфавіті X_i, $i = 1, 2$.

15. Навести приклад такої нетривіальної регулярної події R в алфавіті X ($R \ne X^*$ та $R \ne \varnothing$) і нерегулярної події Q, що подія

(а) $R \cup Q$ — регулярна; (б) подія $R \cup Q$ — нерегулярна.

16. Нехай R — регулярна подія, а S — скінченна подія. Довести регулярність події:

(а) $R \cup S$; (б) $R \setminus S$.

17. Нехай Q — нерегулярна подія, а S — скінченна подія. Довести, що подія T нерегулярна:

(а) $T = Q \cup S$; (б) $T = Q \setminus S$.

18. Довести, що алгебра $A = \langle R, \{\cup, \cap, \bar{\ }, \varnothing, X^*\}\rangle$ регулярних подій в алфавіті X є булевою алгеброю .

7.14. Абстрактна та структурна теорії автоматів

У рамках описаної вище теорії ми абстрагуємось як від внутрішньої структури й інтерпретації самого автомата, так і від змістовного трактування його станів, вхідних і вихідних сигналів тощо. Єдиним відступом, зробленим у бік інтерпретації цієї математичної моделі, була згадка про автоматний час. Такий підхід дістав назву ***абстрак-***

тної (або *поведінкової*) ***теорії автоматів***. Не цікавлячись способом реалізації автомата, абстрактна теорія вивчає тільки проблеми, пов'язані з зовнішньою поведінкою автомата, із тим, як він перетворює вхідну інформацію у вихідну та у відповідну послідовність станів автомата, або з тим, які множини слів у певному алфавіті може розпізнати (відрізнити) такий автомат. Досліджувану таким чином модель називають також ***абстрактним*** (скінченним) ***автоматом***. Абстрактна теорія по суті є частиною (розділом) загальної теорії алгоритмів (у цьому разі — алгоритмів зі скінченною пам'яттю), більшість властивостей яких можна вивчати (досліджувати) безвідносно до способу реалізації.

На відміну від абстрактної теорії автоматів у так званій ***структурній теорії*** роблять подальші кроки в напрямі змістовної інтерпретації та врахування більшого числа властивостей реальних дискретних пристроїв. Основна відмітна особливість структурної теорії автоматів полягає в тім, що вона враховує структуру вхідних і вихідних сигналів автомата, а також його внутрішню структуру. З погляду структурної теорії автомат — це композиція (структура, схема, з'єднання, мережа), складена з простіших (елементарних) автоматів. Отриманий таким способом автомат називають ***структурним*** (скінченним) ***автоматом*** (*автоматною*, або *логічною*, *схемою*).

Абстрактна і структурна теорії автоматів — це два розділи загальної теорії автоматів, що взаємопов'язані, взаємодіють і доповнюють один одного. На рівні абстрактної теорії простіше, природніше й ефективніше можна розв'язувати задачі визначення об'єму пам'яті (складності) автомата, рівносильних перетворень і оптимізації алгоритму його роботи, розробки різноманітних мов для опису поведінки автоматів і порівняння можливостей та дослідження властивостей цих мов.

Основні задачі структурної теорії автоматів — це також задачі аналізу, синтезу й оптимізації (мінімізації) схеми автомата. Остання задача полягає в побудові схеми автомата, що містить найменшу можливу кількість елементів. Окрім оптимізації автомата за кількістю елементів до задач структурної теорії автоматів належать забезпечення надійності, перешкодостійкості, економічності, швидкодії схем, тобто задачі оптимізації схем за різними технічними параметрами і показниками. У структурній теорії автоматів зазвичай розглядають лише математичні формальні методи розв'язання названих технічних задач.

Центральною в структурній теорії автоматів, безумовно, є задача синтезу, тобто побудова схеми (композиції елементарних автоматів), що реалізує задану поведінку. Початковою інформацією для структурного синтезу зазвичай служить опис автомата в термінах абстрактної теорії. Цю ситуацію доречно описати в алгебричній термінології. У структурній теорії ми по суті маємо справу з алгеброю, носієм якої є множина елементарних автоматів, а сигнатура складається з точно означених операцій композиції для елементарних автоматів (останні, до речі, задають за допомогою вищеозначених моделей абстрактної теорії). Ця алгебра подібна до алгебри комбінаційних схем (див. розд. 6.14), із тією відмінністю, що алгебра структурних автоматів має справу з істотно складнішими елементами, а також використовує складніші композиції.

7.15. Методи композиції автоматів

До основних способів композиції автоматів відносять паралельне і послідовне сполучення, а також сполучення зі зворотним зв'язком.

1. ***Паралельне сполучення***. Результатом паралельного сполучення автоматів $A_1 = (X, Y_1, U_1, \delta_1, \lambda_1)$ і $A_2 = (X, Y_2, U_2, \delta_2, \lambda_2)$ називатимемо такий автомат $A = (X, Y, U, \delta, \lambda)$:

- множина станів — $U = U_1 \times U_2$;
- вихідний алфавіт — $Y = \varphi\,(Y_1 \times Y_2)$;
- функцію переходів δ: $U \times X \to U$ означено так: якщо $a_m = (a_{m_1}, a_{m_2}) \in U$, то $\delta(a_m, x_k) = (\delta_1(a_{m_1}, x_k), \delta_2(a_{m_2}, x_k))$;
- функцію виходів λ: $U \times X \to Y$ означено так:

$$\lambda(a_m, x_k) = \varphi(\lambda_1(a_{m_1}, x_k), \lambda_2(a_{m_2}, x_k));$$

- φ: $Y_1 \times Y_2 \to Y$ — деяке відображення.

Паралельне сполучення автоматів A_1 і A_2 зручно зображати за допомогою такої діаграми (рис. 7.18).

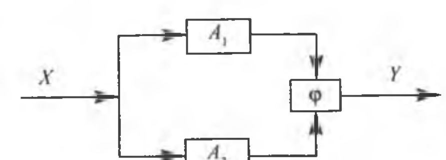


Рис. 7.18

Тут розглянуто паралельне сполучення автоматів, що функціонують у спільному вхідному алфавіті. Подібним чином можна означити паралельне сполучення автоматів, які мають відмінні вхідні алфавіти. Неважко узагальнити введену операцію на довільну кількість елементарних автоматів, а також означити її для моделі Мура.

2. ***Послідовне сполучення***. Результатом послідовного сполучення двох автоматів $A_1 = (X, Y_1, U_1, \delta_1, \lambda_1)$ і $A_2 = (X_2, Y_2, U_2, \delta_2, \lambda_2)$ за умови $X_2 = Y_1$ назвемо такий автомат $A = (X, Y_2, U, \delta, \lambda)$:

- множина станів — $U = U_1 \times U_2$;
- функцію переходів δ: $X \times U \to U$ означено так: якщо $a_m = (a_{m_1}, a_{m_2}) \in U$, то $\delta(a_m, x_k) = (\delta_1(a_{m_1}, x_k), \delta_2(a_{m_2}, \lambda_1(a_{m_1}, x_k)))$;
- функцію виходів λ: $X \times U \to Y$ задано співвідношенням

$$\lambda(a_m, x_k) = \lambda_2(a_{m_2}, \lambda_1(a_{m_1}, x_k)).$$

Відповідну діаграму подано на рис. 7.19.

Рис. 7.19

3. ***Сполучення зі зворотним зв'язком.*** Результатом сполучення зі зворотним зв'язком автоматів $A_1 = (X_1, Y_1, U_1, \delta_1, \lambda_1)$ і $A_2 = (X_2, Y_2, U_2, \delta_2, \mu)$, де A_2 — автомат Мура і $X_2 = Y_1$, назвемо такий автомат $A = (X, Y, U, \delta, \lambda)$:

- множина станів — $U = U_1 \times U_2$;
- вхідний і вихідний алфавіти — $X = X_1$ і $Y = Y_2$;
- функцію переходів δ: $X \times U \to U$ означено так:

якщо $a_m = (a_{m_1}, a_{m_2}) \in U$, то

$$\delta(a_m, x_k) = (\delta_1(a_{m_1}, \gamma(x_k, \mu(a_{m_2}))), \delta_2(a_{m_2}, \lambda_1(a_{m_1}, \gamma(x_k, \mu(a_{m_2})))));$$

- функція виходів λ: $X \times U \to Y$ має вигляд

$$\lambda(a_m, x_k) = \lambda_1(a_{m_1}, \gamma(x_k, \mu(a_{m_2})));$$

- γ: $X_1 \times Y_2 \to X_1$ — якесь алфавітне відображення.

Діаграму, що відповідає сполученню зі зворотним зв'язком двох автоматів, подано на рис. 7.20.

Застосовуючи до елементарних автоматів описані вище методи композиції, можна отримувати як завгодно складні та різноманітні схеми. Власне композиція автоматів полягає в ототожненні (в абстрактному теоретичному трактуванні) або з'єднанні (у практичному технічному розумінні) елементарних вхідних і вихідних каналів (сиг-

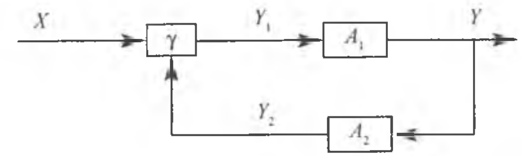


Рис. 7.20

налів) об'єднуваних у єдину схему елементарних автоматів. Для того щоб будь-яка з операцій композиції була завжди виконуваною й коректною, необхідно, щоб усі елементарні автомати мали той самий спільний алфавіт (вхідний і вихідний). Такий *алфавіт* називають ***структурним***. На практиці структурним алфавітом найчастіше є двійковий алфавіт {0, 1}.

Мають бути виконані також *умови коректної побудови схеми* структурного автомата (або умови коректності виконання операцій композиції в алгебрі структурних автоматів) [8]. Неформально основні умови коректності схеми можна звести до таких:

- у будь-який момент автоматного часу в усіх каналах схеми є сигнал;
- неприпустима неоднозначність сигналів у каналі, тобто не можна ототожнювати декілька вихідних сигналів з одним вхідним;
- зворотні зв'язки у схемі мають бути побудовані коректно.

Розглянемо метод розв'язання *проблеми аналізу* у структурній теорії автоматів. Якщо S — схема, складена з автоматів $A_1, A_2, \ldots, A_n$ за допомогою розглянутих вище операцій композиції, то неважко побудувати відповідний абстрактний автомат A, що описує поведінку схеми S мовою автоматних відображень.

Виділимо у схемі S вхідні канали $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ автоматів A_i, $i = 1, 2, \ldots, n$, які не було ототожнено з жодним вихідним каналом. Природно називати такі *вхідні канали зовнішніми*. Аналогічно виділимо *зовнішні вихідні канали* $y_{j_1}, y_{j_2}, \ldots, y_{j_p}$. Множиною станів автомата A є сукупність векторів $(a_{l_1}, a_{l_2}, \ldots, a_{l_n})$ станів автоматів $A_1, A_2, \ldots, A_n$ відповідно. Поставимо у взаємно однозначну відповідність кожному з векторів значень вхідних і вихідних каналів схеми деякий символ з абстрактних алфавітів X і Y відповідно. Так само кожному з векторів станів $(a_{l_1}, a_{l_2}, \ldots, a_{l_n})$ поставимо у взаємно однозначну відповідність деякий символ (абстрактний стан) з алфавіту U. Функцію переходів δ і функцію виходів λ шуканого абстрактного автомата $(X, Y, U, \delta, \lambda)$ можна визначити послідовно за допомогою оператора суперпозиції й описаних вище правил композиції для функцій $\delta_1, \delta_2, \ldots, \delta_n$ і $\lambda_1, \lambda_2, \ldots, \lambda_n$ елементарних автоматів. Із цієї побудови, зокрема, випливає, що число станів автомата A не перевищує добутку кількостей станів усіх автоматів $A_1, A_2, \ldots, A_n$.

7.13. Задачі і вправи

1. Побудувати таблиці, граф і матриці переходів і виходів автомата, який є паралельним сполученням автоматів A_1 і A_2, заданих суміщеними таблицями переходів/виходів.

A_1

δ_1/λ_1	a_1	a_2	a_3
x_1	$a_2/0$	$a_3/1$	$a_3/1$
x_2	$a_1/0$	$a_1/0$	$a_2/0$

A_2

δ_2/λ_2	b_1	b_2
x_1	$b_2/1$	$b_2/0$
x_2	$b_1/1$	$b_2/1$

Функціональне перетворення φ означено співвідношенням $\varphi((y_i, y_j)) = \nu((y_i, y_j))$, де $\nu(\alpha)$ — номер двійкового кортежу α (див. розд. 6.1).

2. Означити паралельне сполучення двох автоматів Мура.

3. Означити паралельне сполучення n автоматів Мілі $A_k = (X_k, Y_k, U_k, \delta_k, \lambda_k)$, $k = 1, 2, \ldots, n$.

4. Довести, що для будь-яких регулярних подій R_1 і R_2 відповідно у скінченних алфавітах X_1 і X_2 декартів добуток $R_1 \times R_2$ є регулярною подією в алфавіті $X = X_1 \times X_2$.

5. Побудувати таблиці, граф і матриці переходів і виходів автомата, що є результатом послідовного сполучення автоматів A_1 і A_2, заданих суміщеними таблицями переходів/виходів.

A_1

δ_1/λ_1	a_1	a_2
x_1	$a_2/0$	$a_1/1$
x_2	$a_1/1$	$a_2/1$
x_3	$a_2/0$	$a_2/0$

A_2

δ_2/λ_2	b_1	b_2	b_3
0	b_2/y_1	b_2/y_3	b_3/y_2
1	b_1/y_3	b_3/y_2	b_3/y_1

6. Визначити автомат, що є результатом сполучення зі зворотним зв'язком автоматів A_1 і A_2, заданих таблицями.

A_1

δ_1/λ_1	a_1	a_2
x_1	$a_1/0$	$a_2/0$
x_2	$a_1/1$	$a_1/1$
x_3	$a_2/0$	$a_2/0$

A_2

μ	y_1	y_3	y_2
δ_2	b_1	b_2	b_3
0	b_2	b_1	b_3
1	b_1	b_3	b_1

Функціональне перетворення γ задано таблицею.

γ	y_1	y_2	y_3
x_1	0	1	1
x_2	1	1	0
x_3	0	0	1

7. Розв'язати задачу аналізу для структурних автоматів, побудованих у задачах 1, 5 і 6.

7.16. Канонічний метод структурного синтезу автоматів

Як було зазначено вище, основною задачею структурної теорії автоматів є задача знаходження загальних прийомів побудови (синтезу) структурних схем автоматів на основі композиції автоматів, що належать до наперед заданого скінченного набору типів автоматів. Точніше, ***задачу структурного синтезу*** автоматів можна сформулювати так.

Нехай є якась скінченна множина елементарних автоматів у одному й тому самому структурному алфавіті C (припускаємо, що кожен з елементарних автоматів є у цій множині в необмеженій кількості). Нехай задано також довільний абстрактний скінченний автомат A. Потрібно знайти алгоритм, який даватиме змогу за заданим автоматом A будувати деяку композицію елементарних автоматів так, що отриманий у результаті композиції структурний автомат (схема) реалізує відображення, що продовжує відображення, індуковане автоматом A.

Слід зазначити, що далеко не для кожного набору елементарних автоматів ця задача має розв'язок. Тоді ж, коли вона має розв'язок (для довільного скінченного автомата A), будемо говорити, що задана система елементарних автоматів є *структурно повною* (або є *базисом*).

Є багато різних структурно повних наборів елементарних автоматів. Однак на сьогодні немає скільки-небудь ефективних засобів (істотно простіших, ніж метод перебору всіх можливих варіантів) для розв'язання основної задачі структурного синтезу в разі довільного вибору структурно повної системи елементарних автоматів. Тому розглянемо так званий канонічний метод структурного синтезу автоматів, який застосовують для структурно повних систем елементарних автоматів спеціального виду.

Канонічний метод структурного синтезу оперує з елементарними автоматами двох типів. До першого типу належать *автомати з пам'яттю* (*елементи пам'яті*, або *запам'ятовувальні елементи*). Такі автомати мають більше ніж один внутрішній стан. Другу групу становлять *автомати без пам'яті* (тобто автомати з одним внутрішнім станом — *комбінаційні*, або *логічні, елементи*). Результатом композиції логічних елементів завжди є автомат без пам'яті або комбінаційна схема. Кожну комбінаційну схему можна охарактеризувати функцією виходів $y(t) = \lambda(x(t))$, де $x(t)$ — структурний вхідний сигнал, $y(t)$ — структурний вихідний сигнал у той самий момент часу t.

Основна ідея канонічного методу структурного синтезу полягає в тому, щоб звести задачу структурного синтезу довільного автомата до задачі структурного синтезу деякої комбінаційної схеми, тобто схеми автомата без пам'яті. Щоб реалізувати цю ідею, потрібно спеціальним чином відібрати до базису елементи пам'яті. А саме, при канонічному методі структурного синтезу автоматами з пам'яттю, що входять до базису, обирають автомати Мура, які мають властивості повноти систем переходів і виходів. Означимо ці властивості.

Повнота системи переходів у автоматі $A = (X, Y, U, \delta, \mu)$ означає, що для будь-якої впорядкованої пари його станів є вхідний сигнал, що переводить автомат із першого стану у другий. Інакше кажучи, для повноти системи переходів необхідно й достатньо, щоб для довільної пари (a, b) його станів існував вхідний сигнал x, для якого $b = \delta(a, x)$.

Повнота системи виходів в автоматі Мура означає, що кожному стану автомата відповідає свій власний вихідний сигнал, відмінний від вихідного сигналу, що відповідає будь-якому іншому стану.

Інакше кажучи, в автоматі Мура з повною системою виходів функція відміток μ реалізує взаємно однозначну відповідність між множиною станів і множиною його вихідних сигналів. Із цього означення безпосередньо випливає, що в автоматі Мура з повною системою виходів можна ототожнити внутрішні стани та відповідні вихідні сигнали автомата. Отже, операції кодування станів і кодування вихідних сигналів у таких автоматах можна об'єднати, тому стани та відповідні їм вихідні сигнали можуть набути однакових значень кодів у структурному алфавіті. Це дає змогу використовувати той самий структурний алфавіт для кодування як вхідних і вихідних сигналів, так і внутрішніх станів автоматів.

Для автоматів з повною системою переходів уведемо таке означення.

Функцією входів π автомата $A = (X, Y, U, \delta, \mu)$ з повною системою переходів назвемо функцію, яка довільній парі станів $a, b \in U$ автомата A ставить у відповідність множину вхідних сигналів $x \in X$ таких, що $b = \delta(a, x)$.

У таблиці функції входів автомата A рядкам і стовпчикам відповідають його стани. На перетині i-го рядка, що відповідає стану a_i, і j-го стовпчика, що відповідає стану a_j, записують множину $\pi(a_i, a_j)$ вхідних сигналів, що зумовлюють перехід автомата A зі стану a_i в стан a_j.

Теоретичним обґрунтуванням канонічного методу структурного синтезу є така теорема.

Теорема 7.14 (***про структурну повноту***). Будь-яка система елементарних автоматів, яка містить автомат Мура з нетривіальною пам'яттю, повною системою переходів і повною системою виходів, а також будь-яку функціонально повну систему логічних елементів, є структурно повною системою. Існує загальний конструктивний спосіб (*канонічний метод структурного синтезу*), що дає змогу в цьому разі звести задачу структурного синтезу довільного скінченного автомата до задачі структурного синтезу певних комбінаційних схем.

Доведення. Нехай задано довільний абстрактний скінченний автомат $A = (X, Y, U, \delta, \lambda)$. Візьмемо p елементарних автоматів Мура з пам'яттю $A_1, A_2, \ldots, A_p$, що мають повні системи переходів і виходів. Кількість потрібних автоматів визначає співвідношення $|U| \le |U_1| \cdot |U_2| \cdot \ldots \cdot |U_p|$, де U_i — множина станів автомата A_i.

Закодуємо стани автомата A, зіставивши кожному абстрактному стану $\boldsymbol{a}_i$ автомата A впорядкований набір $(a_{i_1}, a_{i_2}, \ldots, a_{i_p})$ станів автоматів A_j, $j = 1, 2, \ldots, p$. Використовуючи властивість повноти системи виходів, коди станів $(a_{i_1}, a_{i_2}, \ldots, a_{i_p})$ можна ототожнити з кодами вихідних сигналів $(v_{i_1}, v_{i_2}, \ldots, v_{i_p})$ автоматів $A_1, A_2, \ldots, A_p$, де v_{i_j} — вихідний сигнал, що відповідає в автоматі A_j стану a_{ij}. Позначимо структурний вихідний сигнал (чи, що те саме, структурний стан) $(v_{i_1}, v_{i_2}, \ldots, v_{i_p})$ структурного автомата через $\boldsymbol{v}_i$.

Закодуємо вхідні сигнали заданого автомата A, зіставивши абстрактному сигналу $\boldsymbol{x}_i$ структурний вхідний сигнал $(x_{i_1}, x_{i_2}, \ldots, x_{i_p})$.

Тоді функцію переходів і закон функціонування автомата A можна записати у вигляді $\boldsymbol{a}(t+1) = \delta(\boldsymbol{a}(t), \boldsymbol{x}(t))$ або (з урахуванням взаємно однозначної відповідності станів і вихідних сигналів) у вигляді

$$\boldsymbol{v}(t+1) = \delta(\boldsymbol{v}(t), \boldsymbol{x}(t)). \qquad (7.13)$$

Перехід автомата A зі стану $\boldsymbol{v}(t)$ в стан $\boldsymbol{v}(t+1)$ складається з відповідних переходів в елементарних автоматах $A_1, A_2, \ldots, A_p$. Для забезпечення переходу автомата A_i зі стану $v_i(t)$ в стан $v_i(t+1)$ відповідно до заданої функції переходів δ потрібно в момент часу t подати на вхід елементарного автомата A_i певний вхідний сигнал $u_i(t)$, $i = 1, 2, \ldots, p$. Виходячи з властивості повноти системи переходів усіх елементарних автоматів A_i, покладемо $u_i(t) = \pi_i(v_i(t), v_i(t+1))$, де π_i — функція входів автомата A_i, $i = 1, 2, \ldots, p$, а для $u_i(t)$ обрано один з елементів її множини значень.

Природним чином отримаємо структурний вхідний сигнал пам'яті схеми S автомата A:

$$\boldsymbol{u}(t) = \pi(\boldsymbol{v}(t), \boldsymbol{v}(t+1)). \qquad (7.14)$$

Підставивши у співвідношення (7.14) значення $\boldsymbol{v}(t+1)$ з формули (7.13), одержимо

$$\boldsymbol{u}(t) = \pi(\boldsymbol{v}(t), \delta(\boldsymbol{v}(t), \boldsymbol{x}(t))) = \rho(\boldsymbol{v}(t), \boldsymbol{x}(t)). \qquad (7.15)$$

Задана цією формулою функція $\rho(\boldsymbol{v}, \boldsymbol{x})$ називається (векторною) ***функцією збудження*** схеми S автомата A, або *функцією входів пам'яті*. Вона визначає, який структурний вхідний сигнал $\boldsymbol{u}(t)$ в будь-який момент часу t потрібно подати на входи елементів пам'яті схеми S автомата A, що перебуває у стані $\boldsymbol{v}(t)$, якщо зовнішнім структурним вхідним сигналом є $\boldsymbol{x}(t)$, а у наступний момент часу $t+1$ схема S має опинитися в стані $\boldsymbol{v}(t+1)$, визначеному заданою функцією переходів δ.

Завдяки збігу всіх сигналів у співвідношенні (7.15) за часом сигнал $\boldsymbol{u}(t)$ можна розглядати як вихідний сигнал якоїсь комбінаційної схеми, на вхід якої подано об'єднаний структурний вхідний сигнал $(\boldsymbol{v}(t), \boldsymbol{x}(t))$. Реалізувавши цю комбінаційну схему за допомогою наявних логічних елементів із функціонально повної системи, ми забезпечимо тим самим можливість реалізації всіх переходів зі стану в стан, передбачених функцією переходів δ заданого автомата A.

Побудовану таким чином комбінаційну схему називають *схемою зворотних зв'язків*, а рівняння, що виражають залежність сигналів, які подають на входи запам'ятовувальних елементів, і вихідних сигналів схеми від сигналів, які надходять на вхід схеми (автомата) в цілому, і сигналів, що знімають з виходів елементів пам'яті, називають ***канонічними*** (векторними) ***рівняннями*** схеми S автомата A. Перше з канонічних рівнянь має вигляд

$$\boldsymbol{u} = \rho(\boldsymbol{v}, \boldsymbol{x}). \qquad (7.16)$$

Зауважимо, що функція збудження $\rho(\boldsymbol{v}, \boldsymbol{x})$ схеми автомата визначається, взагалі кажучи, неоднозначно, бо неоднозначним у загальному випадку є вибір значення функції входів π_i елементарних автоматів $A_1, A_2, \ldots, A_p$. Цим можна скористатися для одержання простішої схеми. Інше джерело оптимізації синтезованої схеми — вибір того чи іншого способу кодування станів (так звана *проблема кодування*).

І нарешті, розглянемо питання про реалізацію шуканою схемою функції виходів автомата A. Залежно від того, чи є заданий автомат A автоматом Мілі або автоматом Мура, утворення його вихідних сигналів буде визначатися або законом $\boldsymbol{y}(t+1) = \lambda(\boldsymbol{v}(t), \boldsymbol{x}(t))$, або законом $\boldsymbol{y}(t) = \mu(\boldsymbol{v}(t))$, що задає друге канонічне рівняння схеми S. В обох випадках потрібний структурний вихідний сигнал реалізує комбінаційна схема, яку називають *схемою виходів*. Для автомата Мілі ця схема реалізує векторну функцію $\boldsymbol{y} = \lambda(\boldsymbol{v}, \boldsymbol{x})$, а для автомата Мура — векторну функцію $\boldsymbol{y} = \mu(\boldsymbol{v})$. Для обох моделей автоматів можливість реалізації схеми виходів випливає з функціональної повноти системи логічних елементів. Теорему доведено.

Структурна схема будь-якого автомата, синтезованого відповідно до канонічного методу структурного синтезу, має вигляд, поданий на рис. 7.21. У канонічному методі структурного синтезу традиційно зображають структурну схему автомата у вигляді двох частин: набору елементарних автоматів з пам'яттю, що складають *пам'ять* (або *запам'ятову-*

вальну частину), і двох комбінаційних схем — схеми виходів і схеми зворотних зв'язків, які зазвичай об'єднують в одну *комбінаційну частину* (рис. 7.21).

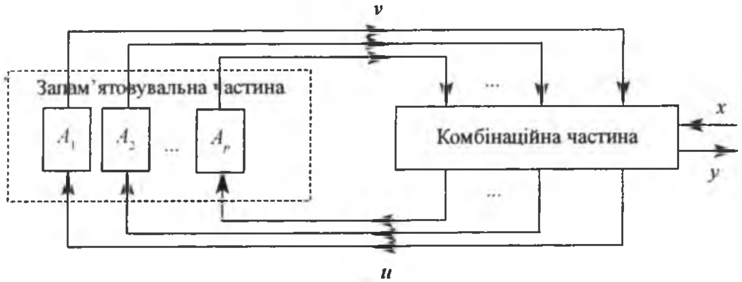


Рис. 7.21

На закінчення ще раз підкреслимо, у чому полягає основна перевага канонічного методу структурного синтезу: проблема синтезу довільного скінченного автомата зводиться до проблеми синтезу відповідних комбінаційних схем — схеми зворотних зв'язків для ρ та схеми виходів для λ. Наведемо загальну схему реалізації цього методу.

Основні етапи структурного синтезу

1. Обрати структурний алфавіт C.
2. Поставити у відповідність кожному символу $x_i \in X$ вхідного алфавіту абстрактного автомата A сигнал $\boldsymbol{x}_c$ у структурному алфавіті C, тобто здійснити *кодування вхідних сигналів*.
3. Аналогічно виконати *кодування абстрактних вихідних сигналів* $y_i \in Y$ у тому самому структурному алфавіті C.
4. Вибрати елементарний автомат, що задовольняє умови повноти системи переходів і повноти системи виходів. Визначити потрібну кількість елементарних автоматів обраного типу.
5. *Здійснити кодування станів*, тобто поставити у відповідність кожному абстрактному стану $a_k \in U$ набір станів $\boldsymbol{a}_c$ елементарних автоматів.
6. Визначити логічні (перемикальні) функції: функцію збудження $\boldsymbol{u} = \rho(\boldsymbol{v}, \boldsymbol{x})$ і функцію виходів $\boldsymbol{y} = \lambda(\boldsymbol{v}, \boldsymbol{x})$ (або $\boldsymbol{y} = \mu(\boldsymbol{v})$).

Розглянемо приклад застосування канонічного методу структурного синтезу.

Приклад 7.9. Нехай задано абстрактний автомат Мілі $A = (X, Y, U, \delta, \lambda)$ за допомогою його таблиць переходів і виходів

δ	a_1	a_2	a_3	a_4
x_1	a_2	a_3	a_4	a_1
x_2	a_1	a_4	a_1	a_1
x_3	a_4	a_1	a_1	a_1

λ	a_1	a_2	a_3	a_4
x_1	y_1	y_2	y_3	y_4
x_2	y_2	y_3	y_5	y_1
x_3	y_4	y_3	y_2	y_2

1. Оберемо структурний алфавіт $C = \{1, 2, 3\}$.
2. Виконаємо кодування абстрактних вхідних сигналів, поставивши у відповідність символу x_i вхідного алфавіту автомата A сигнал (код довжини 1) i в структурному алфавіті C, $i = 1, 2, 3$.
3. Закодуємо у структурному алфавіті C вихідні сигнали: символу y_1 поставимо у відповідність код (1, 1), символу y_2 — код (1, 2), символу y_3 — (2, 1), символу y_4 — (2, 2), символу y_5 — (3, 1).
4. Елементарним автоматом Мура з пам'яттю та повною системою переходів і повною системою виходів оберемо автомат Z, заданий відміченою таблицею переходів

μ	1	2	3
δ	*a*	*b*	*c*
1	*a*	*c*	*b*
2	*b*	*a*	*c*
3	*c*	*b*	*a*

Враховуючи взаємно однозначну відповідність між множиною станів і множиною вихідних сигналів цього автомата, ототожнимо їх: 1 — *a*, 2 — *b*, 3 — *c*. Тоді таблиця переходів елементарного автомата Z набере вигляду

δ	1	2	3
1	1	3	2
2	2	1	3
3	3	2	1

Кількість елементарних автоматів Z, потрібних для реалізації даного автомата A, дорівнює 2, бо $4 < 3 \cdot 3$.

5. Виконаємо кодування станів автомата A: стану a_1 поставимо у відповідність код (1, 2), стану a_2 — код (3, 1), стану a_3 — код (2, 2), стану a_4 — код (2, 3). Зауважимо, що обраний варіант кодування є лише одним із багатьох можливих. Для зручності укладемо так звані кодовані таблицю переходів і таблицю виходів автомата A.

δ	(1, 2)	(3, 1)	(2, 2)	(2, 3)
1	(3, 1)	(2, 2)	(2, 3)	(1, 2)
2	(2, 2)	(2, 3)	(1, 2)	(1, 2)
3	(2, 3)	(1, 2)	(1, 2)	(1, 2)

λ	(1, 2)	(3, 1)	(2, 2)	(2, 3)
1	(1, 1)	(1, 2)	(2, 1)	(2, 2)
2	(1, 2)	(2, 1)	(3, 1)	(1, 1)
3	(2, 2)	(2, 1)	(1, 2)	(1, 2

Запишемо таблицю функції входів елементарного автомата Z, тобто таблицю функції π:

π	1	2	3
1	1	2	3
2	2	3	1
3	3	1	2

6. Перейдемо до визначення функції збудження і функції виходів шуканої схеми, тобто до укладання таблиць цих функцій (рис. 7.22). Опишемо процедуру заповнення першого рядка цієї таблиці.

Якщо стан (v_1, v_2) структурного автомата в момент часу t дорівнює (1, 2) (другий і третій стовпчики таблиці), а на його вхід подано структурний сигнал 1 (перший стовпчик таблиці), то згідно з кодованою таблицею переходів для функції δ в момент $t + 1$ він має опинитися у стані (3, 1) (останні два стовпчики таблиці). Отже, перший елементарний автомат Z_1 має змінити свій стан з 1 на 3, а другий Z_2 — з 2 на 1. Перший із цих переходів буде забезпечено, якщо на вхід Z_1 подати сигнал $u_1(t) = \pi(1, 3)$, який дорівнює 3 (див. таблицю функції входів π). А перехід для Z_2 буде виконано під дією сигналу $u_2(t) = \pi(2, 1)$, що дорівнює 2. Водночас при такій комбінації "стан—вхідний сигнал" на виході схеми слід забезпечити сигнал (1, 1), тобто $y_1(t) = 1$ і $y_2(t) = 1$.

Таким чином, три перші та два останні стовпчики таблиці слід заповнювати за допомогою кодованої таблиці переходів. Четвертий і п'ятий стовпчики отримаємо, виходячи зі співвідношень

$u_1(t) = \pi(v_1(t), v_1(t + 1))$ і $u_2(t) = \pi(v_2(t), v_2(t + 1))$. Значення для $y_1(t)$ й $y_2(t)$ можна визначити за допомогою кодованої таблиці виходів: $\big(y_1(t), y_2(t)\big) = \lambda\big(\big(v_1(t), v_2(t)\big), x(t)\big)$.

$x(t)$	$v_1(t)$	$v_2(t)$	$u_1(t)$	$u_2(t)$	$y_1(t)$	$y_2(t)$	$v_1(t+1)$	$v_2(t+1)$
1	1	2	3	2	1	1	3	1
1	3	1	1	2	1	2	2	2
1	2	2	3	1	2	1	2	3
1	2	3	2	1	2	2	1	2
2	1	2	2	3	1	2	2	2
2	3	1	1	3	2	1	2	3
2	2	2	2	3	3	1	1	2
2	2	3	2	1	1	1	1	2
3	1	2	2	1	2	2	2	3
3	3	1	3	2	2	1	1	2
3	2	2	2	3	1	2	1	2
3	2	3	2	1	1	2	1	2

Рис. 7.22

Отже, слід побудувати дві окремі комбінаційні схеми (зворотних зв'язків і виходів), кожну з трьома вхідними каналами x, v_1 і v_2 та двома вихідними каналами, відповідно u_1, u_2 та y_1, y_2, або одну об'єднану схему з трьома вхідними x, v_1 і v_2 та чотирма вихідними каналами u_1, u_2, y_1, y_2. Для практичної реалізації такої схеми потрібна функціонально повна система базових комбінаційних елементів, що реалізують перемикальні функції в алфавіті $C = \{1, 2, 3\}$, тобто функції типу $C^n \to C$.

Алфавіт $C = \{1, 2, 3\}$ було обрано для того, щоб продемонструвати можливість побудови схеми в довільному структурному алфавіті. Однак на практиці зазвичай структурним алфавітом є двійковий алфавіт $B = \{0, 1\}$. Розглянемо ще один приклад структурного синтезу схеми для випадку двійкового алфавіту.

Приклад 7.10. Нехай задано скінченний частковий автомат Мілі A за допомогою його таблиць переходів і виходів.

δ	a_1	a_2	a_3	a_4
x_1	a_1	a_4	a_4	–
x_2	–	a_3	a_1	–
x_3	a_2	–	a_2	a_3

λ	a_1	a_2	a_3	a_4
x_1	y_2	y_1	y_3	–
x_2	–	y_3	y_1	–
x_3	y_1	–	y_4	y_2

1. Структурний алфавіт — $C = \{0, 1\}$.

2. Виконаємо кодування абстрактних вхідних сигналів, поставивши у відповідність сигналу x_1 вхідного алфавіту автомата A сигнал (код) (0, 0) у структурному алфавіті C, x_2 — код (1, 0), x_3 — код (0, 1).

3. Закодуємо у структурному алфавіті C вихідні сигнали: символу y_1 поставимо у відповідність код (0, 0), y_2 — (1, 0), y_3 — (0, 1), y_4 — (1, 1).

4. Елементарним автоматом Мура з пам'яттю та повною системою переходів і повною системою виходів оберемо автомат Z, заданий відміченою таблицею переходів

μ	0	1
δ	a	b
0	a	b
1	b	a

З урахуванням взаємно однозначної відповідності між множи-ною станів і множиною вихідних сигналів зробимо такі ототожнення: 0 — a, 1 — b. Тоді таблиця переходів елементарного автомата Z набере вигляду

δ	0	1
0	0	1
1	1	0

Кількість елементарних автоматів Z, потрібних для реалізації даного автомата A, дорівнює 2, бо $4 = 2 \cdot 2$.

5. Закодуємо стани автомата A: стану a_1 поставимо у відповідність код (0, 0), стану a_2 — код (0, 1), стану a_3 — код (1, 0), стану a_4 — код (1, 1). Кодовані таблиці переходів і виходів автомата A матимуть такий вигляд:

δ	00	01	10	11
1	00	11	11	–
2	–	10	00	–
3	01	–	01	10

λ	00	01	10	11
1	10	00	01	–
2	–	01	00	–
3	00	–	11	10

Таблиця функції входів π автомата Z має вигляд

π	0	1
0	0	1
1	1	0

6. Таблиці функції збудження та функції виходів шуканої схеми будуємо як і раніше.

$x_1(t)$	$x_2(t)$	$v_1(t)$	$v_2(t)$	$u_1(t)$	$u_2(t)$	$y_1(t)$	$y_2(t)$	$v_1(t+1)$	$v_2(t+1)$
0	0	0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0	1	1
0	0	1	0	0	1	0	1	1	1
1	0	0	1	1	1	0	1	1	0
1	0	1	0	1	0	0	0	0	0
0	1	0	0	0	1	0	0	0	1
0	1	1	0	1	1	1	1	0	1
0	1	1	1	0	1	1	0	1	0

Запишемо ДДНФ для функцій збудження та функцій виходів:

$$u_1 = \overline{x}_1\overline{x}_2\overline{v}_1v_2 \vee x_1\overline{x}_2\overline{v}_1v_2 \vee x_1\overline{x}_2v_1\overline{v}_2 \vee \overline{x}_1x_2v_1\overline{v}_2;$$

$$u_2 = \overline{x}_1\overline{x}_2v_1\overline{v}_2 \vee x_1\overline{x}_2\overline{v}_1v_2 \vee \overline{x}_1x_2\overline{v}_1\overline{v}_2 \vee \overline{x}_1x_2v_1\overline{v}_2 \vee \overline{x}_1x_2v_1v_2;$$

$$y_1 = \overline{x}_1\overline{x}_2\overline{v}_1\overline{v}_2 \vee \overline{x}_1x_2v_1\overline{v}_2 \vee \overline{x}_1x_2v_1v_2;$$

$$y_2 = \overline{x}_1\overline{x}_2v_1\overline{v}_2 \vee x_1\overline{x}_2\overline{v}_1v_2 \vee \overline{x}_1x_2v_1\overline{v}_2.$$

Після роздільної мінімізації останніх формул отримаємо такі вирази:

$u_1 = x_1 \vee x_2 v_1 \bar{v}_2 \vee \bar{x}_2 v_2;$

$u_2 = x_2 \vee \overline{x_1 v_1} \vee x_1 v_2;$

$y_1 = x_2 v_1 \vee \overline{x_2 v_1 v_2};$

$v_2 = \bar{x}_1 v_1 \bar{v}_2 \vee x_1 v_2.$

Фрагмент комбінаційної схеми, що відповідає функції збудження u_1 зображено на рис. 7.23.

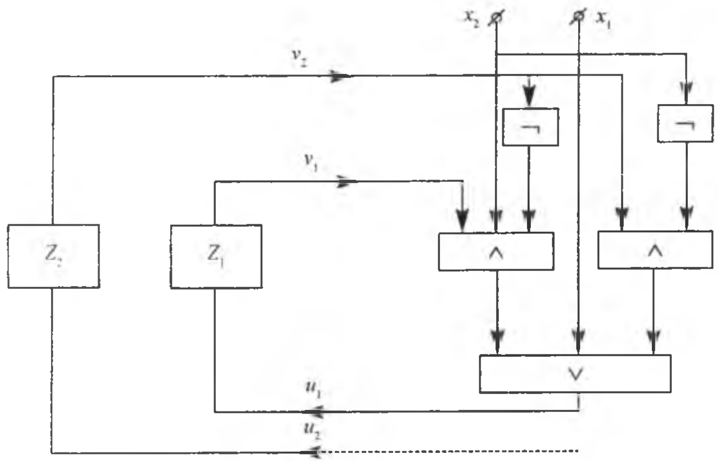


Рис. 7.23

Аналогічно можна накреслити схеми для інших функцій u_2, y_1 та y_2. Зокрема, якщо взяти до уваги те, що друга функція збудження має вигляд $u_2 = y_2 \vee \bar{x}_1 x_2 \overline{v_1 v_2} \vee \bar{x}_1 x_2 v_1 v_2$, то вихід схеми, що відповідає y_2, можна використати для побудови й оптимізації схеми, яка реалізує функцію u_2.

Слід зазначити, що тип елементарного автомата визначається технологією і стандартами виробництва засобів обчислювальної техніки. Для перших комп'ютерів це були різного виду тригери й елементи затримки та комбінаційні елементи для реалізації функціонально повної системи булевих функцій І, АБО та НІ. Згодом, із появою інтегральних схем, "будівельні блоки" стали істотно складнішими та багатофункціональними. Однак основна ідея конструювання схеми (мережі), що реалізує заданий алгоритм функціонування, з множини елементарних автоматів залишилася незмінною.

Практично всі елементарні запам'ятовувальні пристрої, застосовувані в реальних цифрових автоматах, являють собою автомати Мура з повною системою переходів і повною системою виходів. Описану процедуру побудови функцій збудження пам'яті автомата можна істотно спростити, якщо використовувати певні типи елементарних автоматів (тригери затримки, тригери з лічильним або роздільними входами тощо).

Рекомендований алгоритм синтезу структурного автомата можна застосувати лише до скінченних автоматів із порівняно невеликим числом внутрішніх станів. Для складних автоматів попередньо здійснюють їх розбиття на функціональні блоки (*декомпозицію*), кожен з яких синтезують окремо.

Окрема проблема процедури синтезу — *оптимізація* схеми, зокрема за кількістю елементів. Оскільки число елементарних автоматів з пам'яттю визначається числом станів заданого автомата A, то всі зусилля зі зменшення загальної кількості елементів у схемі зосереджують на комбінаційній частині.

Викладені в розд. 6 методи мінімізації формул алгебри логіки та безліч інших методів, що залишилися за межами нашого розгляду, було розроблено саме для спрощення схем обчислювальної техніки. Для кожної з функцій збудження чи виходів ці методи можна застосовувати окремо. Однак якщо в синтезованих комбінаційних схемах припустимі модулі з декількома виходами (так звані *схеми з розгалуженнями*), то такі модулі можна використати для одночасної реалізації кількох булевих функцій, що може значно скоротити загальну кількість елементів. Методи мінімізації, що дозволяють отримати для заданої сукупності булевих функцій формули, в яких окремі підформули збігаються, дістали назву *методів спільної мінімізації* систем булевих функцій. Результат спільної мінімізації, в якому підсумкова кількість потрібних елементів визначається без урахування повторів у спільних підформулах, може виявитися значно оптимальнішим, ніж результат роздільної мінімізації.

7.14. Задачі і вправи

1. Визначити (з точністю до ізоморфізму) всі автомати Мура з двома станами, що мають властивості повноти системи переходів і повноти системи виходів.

2. Побудувати два різні автомати Мура з трьома станами та повними системами переходів і виходів.

3. Побудувати таблиці функцій збудження та функцій виходів структурного автомата, алгоритм роботи якого задано суміщеною таблицею переходів-виходів абстрактного автомата A. Структурний алфавіт $C = \{0, 1, 2\}$, а автомат Мура з повними системами переходів і виходів має таку відмічену таблицю переходів (як і раніше, у таблицях замість сигналів x_i, y_j і станів a_k записуємо відповідні індекси i, j, k):

μ	2	1	3
δ	1	2	3
1	2	2	1
2	1	3	3
3	3	1	2

(а) A

δ/λ	1	2	3	4	5
1	2/1	1/1	3/1	1/2	5/2
2	4/1	1/2	2/3	1/1	3/2
3	5/3	1/3	2/3	3/1	4/2
4	1/2	1/1	3/3	5/3	5/1

(б) A

δ/λ	1	2	3	4	5	6
1	2/1	1/2	1/3	6/4	6/4	4/1
2	3/3	3/3	1/1	2/2	6/4	5/1
3	1/2	2/2	4/1	5/3	6/4	1/1
4	4/3	5/1	4/2	6/3	3/4	2/4
5	1/3	2/2	3/3	4/1	3/1	2/4

4. Побудувати схему структурного автомата, алгоритм роботи якого задано суміщеною таблицею переходів/виходів абстрактного автомата A. Структурний алфавіт $B = \{0, 1\}$, а автомат Мура з повними системами переходів і виходів має таку відмічену таблицю переходів:

μ	2	1
δ	1	2
1	2	1
2	1	2

(а) A

δ/λ	1	2	3	4	5
1	2/1	1/4	3/1	5/2	5/2
2	5/1	2/2	2/3	1/1	3/2
3	5/3	1/4	3/3	3/1	4/4
4	1/2	4/1	3/3	5/3	5/1
5	3/4	2/3	1/1	4/2	2/1

(б) A

δ/λ	1	2	3	4	5	6
1	2/1	1/2	3/2	6/1	6/1	6/1
2	5/2	3/1	1/1	2/2	3/1	5/1
3	6/2	2/2	4/1	5/1	6/2	1/1
4	4/1	5/1	3/2	6/2	3/2	2/2
5	1/1	2/2	3/1	4/1	3/1	2/1

Попередньо застосувавши до виразів, що задають функції збудження та функції виходів комбінаційної частини схеми, методи мінімізації, викладені в розд. 6.16–6.19, намагатися максимально їх спростити.

ВІДПОВІДІ, ВКАЗІВКИ, РОЗВ'ЯЗАННЯ

1. Множини і відношення

1.1. 1. (а), (в). **2.** (а) $\{2, 3, 4, 5, 6\}$; (б) $\{-2, -1, 0, 1, 2\}$; (в) $\{1, 2, 3, 4, 6, 9\}$. **3.** (б), (г), (е), (ж). **4.** (а), (г), (є), (з). **5.** (г), (е). **6.** (б), (в), (е). **7.** (а), (б), (г), (д), (є), (и), (ї). **8.** (а), (б), (д), (е), (є). **9.** $B = \{\varnothing\}$, $A = \varnothing$. **10.** (а) $\{\varnothing, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$; (б) $\{\varnothing, \{\varnothing\}\}$; (в) $\{\varnothing, \{1\}, \{\{2\}\}, \{\{1, 2\}\}, \{1, \{2\}\}, \{1, \{1, 2\}\}, \{\{2\}, \{1, 2\}\}, \{1, \{2\}, \{1, 2\}\}\}$; (г) $\{\varnothing, \{\varnothing\}, \{\{1, 2\}\}, \{\varnothing, \{1, 2\}\}\}$. **11.** (а) $\{\varnothing, \{\varnothing\}, \{\{1\}\}, \{\{2\}\}, \{\{1, 2\}\}, \{\varnothing, \{1\}\}, \{\varnothing, \{2\}\}, \{\varnothing, \{1, 2\}\}, \{\{1\}, \{2\}\}, \{\{1\}, \{1, 2\}\}, \{\{2\}, \{1, 2\}\}, \{\varnothing, \{1\}, \{2\}\}, \{\varnothing, \{1\}, \{1, 2\}\}, \{\varnothing, \{2\}, \{1, 2\}\}, \{\{1\}, \{2\}, \{1, 2\}\}, \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}\}$; (б) $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$.

1.2. 1. (а) $\{1, 2, 3, 5, 6, 7\}$; (б) $\{4, 6\}$; (в) $\varnothing$; (г) $\{1, 2, 3, 5, 6, 7\}$; (д) $\{2, 6, 7\}$; (е) $\varnothing$. **3.** (а) Наприклад, $A = \{1, 2\}$, $B = \{2, 3\}$. Тоді $(A\setminus B)\cup B = \{1, 2, 3\} \neq A$. Доведемо, що $(A\setminus B)\cup B = A$ тоді й тільки тоді, коли $B \subseteq A$. Нехай $(A\setminus B)\cup B = A$. Розглянемо довільний елемент $x\in B$. За властивостями операції об'єднання з $x\in B$ випливає $x\in(A\setminus B)\cup B$. Отже, за умовою $x\in A$. Припустимо, що $B \subseteq A$. Для встановлення рівності $(A\setminus B)\cup B = A$ доведемо такі два включення: $(A\setminus B)\cup B \subseteq A$ і $A \subseteq (A\setminus B)\cup B$. Нехай $x\in(A\setminus B)\cup B$; тоді $((x\in A\setminus B)$ або $(x\in B)) \Rightarrow (((x\in A)$ і $(x\notin B))$ або $(x\in B)) \Rightarrow$ $\Rightarrow ((x\in A)$ або $(x\in A)) \Rightarrow x\in A$. Навпаки, нехай $x\in A$. Тоді $((x\in A)$ і $((x\in B)$ або $(x\notin B))) \Rightarrow (((x\in A)$ і $(x\in B))$ або $((x\in A)$ і $(x\notin B))) \Rightarrow ((x\in B)$ або $(x\in A\setminus B)) \Rightarrow x\in(A\setminus B)\cup B$. (б) Наприклад, $A = \{1, 2\}$, $B = \{2, 3\}$. Тоді $(A\cup B)\setminus B = \{1\} \neq A$. Необхідною й достатньою умовою виконання цієї рівності є $A\cap B = \varnothing$. **4.** (а) $\varnothing$; (б) $\{\varnothing\}$;

(в) $\{\varnothing\}$; (г) $\{\varnothing, \{\varnothing\}\}$; (д) $\{\{\varnothing\}\}$; (е) $\{\varnothing\}$. **5.** (а) Нехай $A \subseteq B$. Доведемо, що $\overline{B} \subseteq \overline{A}$. Розглянемо довільний елемент $x\in\overline{B}$; тоді $x\notin B$. З умови матимемо $x\notin A$, отже $x\in\overline{A}$. Нехай $\overline{B} \subseteq \overline{A}$. Доведемо рівність $A\cup B = B$, тобто два відповідні включення. Для доведення першого включення візьмемо $x\in A\cup B \Rightarrow ((x\in A)$ або $(x\in B)) \Rightarrow$ $\Rightarrow ((x\notin\overline{A})$ або $(x\in B)) \Rightarrow ((x\notin\overline{B})$ або $(x\in B)) \Rightarrow ((x\in B)$ або $(x\in B)) \Rightarrow x\in B$. З іншого боку, якщо $x\in B$, то за властивістю операції об'єднання маємо $x\in A\cup B$. Для доведення того, що з $A\cup B = B$ випливає $A\cap B = A$, візьмемо $x\in A\cap B \Rightarrow ((x\in A)$ і $(x\in B)) \Rightarrow x\in A$. Навпаки, якщо $x\in A$, то $x\in A\cup B$, тому, за умовою $x\in B$. З $x\in A$ і $x\in B$ за означенням операції перетину маємо $x\in A\cap B$. Нехай виконується рівність $A\cap B = A$. Доведемо, що $A\setminus B = \varnothing$. Застосуємо метод доведення від супротивного. Припустімо, що існує елемент $x\in A\setminus B \Rightarrow ((x\in A)$ і $(x\notin B)) \Rightarrow ((x\in A)$ і $(x\notin A\cap B)) \Rightarrow$ $\Rightarrow ((x\in A)$ і $(x\notin A))$. Прийшли до суперечності. Отже припущення, що множина $A\setminus B$ непорожня, неправильне. Нарешті, нехай $A\setminus B = \varnothing$. Доведемо включення $A \subseteq B$. Розглянемо елемент $x\in A$. Якщо припустити, що $x\notin B$, то отримаємо $x\in A\setminus B$, що суперечить умові. Отже, $x\in B$. Таким чином, доведено рівносильність усіх п'яти співвідношень. **6.** (а) $A = B$. Справді, якщо $A = B$, то $A \cup B = A\cap B$. Навпаки, нехай $A\cup B = A\cap B$. Розглянемо $x\in A \Rightarrow x\in A \cup B \Rightarrow x\in A\cap B \Rightarrow x\in B$. Аналогічно, якщо $x\in B \Rightarrow x\in A \cup B \Rightarrow$ $\Rightarrow x\in A\cap B \Rightarrow x\in A$. (б) $A = B$; (в) $A = \overline{B}$; (г) $A = B = \varnothing$; (д) $A\cap B = \varnothing$; (е) $A \subseteq B$; (є) $A = B = \varnothing$; (ж) $A\cap B = \varnothing$. **7.** (а) Доведемо методом від супротивного, що такі множини A, B і C не існують. З умови $A\cap B \neq \varnothing$ випливає, що існує елемент $x\in A\cap B$, тобто $x\in A$ й $x\in B$. Із другої умови маємо $x\notin C$. Отже, $x\in(A\cap B)\setminus C$, що суперечить останній умові. (б) Ні. **8.** Доведемо дві рівності $(A\cap B)\cup A = A$ і $(A\cup B)\cap A = A$. Нехай $x\in(A\cap B)\cup A \Rightarrow (((x\in A)$ і $(x\in B))$ або $(x\in A)) \Rightarrow$ $\Rightarrow ((x\in A)$ або $(x\in A)) \Rightarrow x\in A$. Навпаки, якщо $x\in A$, то за властивістю операції об'єднання $x\in(A\cap B)\cup A$. Для доведення другої рівності розглянемо $x\in(A\cup B)\cap A$. Тоді $((x\in A\cup B)$ і $(x\in A)) \Rightarrow x\in A$. Якщо ж $x\in A$, то $x\in A\cup B$. Отже, $x\in(A\cup B)\cap A$. **9.** Такими умовами є $A \subseteq B$ або $B \subseteq A$. Справді, якщо виконується одна з цих умов (наприклад, $A \subseteq B$), то $\beta(A) \subseteq \beta(B)$, $\beta(A)\cup\beta(B) = \beta(B)$, $A\cup B = B$ і

$\beta(A \cup B) = \beta(B)$. Навпаки, якщо $\beta(A) \cup \beta(B) = \beta(A \cup B)$, то для довільної підмножини $C \subseteq A \cup B$ маємо $C \in \beta(A \cup B) \Rightarrow C \in \beta(A) \cup \beta(B) \Rightarrow$ $\Rightarrow ((C \in \beta(A))$ або $(C \in \beta(B))) \Rightarrow ((C \subseteq A)$ або $(C \subseteq B))$. Звідси випливає, що або $A \cup B \subseteq A$, або $A \cup B \subseteq B$, тобто або $A \cup B = A$, або $A \cup B = B$. Останні співвідношення рівносильні відповідно $B \subseteq A$ або $A \subseteq B$ (див. задачу 1.2.5(а)). **10.** (а) $B = \varnothing$. Якщо $B = \varnothing$, то $A \Delta B = A$. Навпаки, нехай $A \Delta B = A$ і припустімо, що $B \neq \varnothing$. Візьмемо $x \in B$ та розглянемо два випадки: якщо $x \in A$, то $x \notin A \backslash B$ і $x \notin B \backslash A$, отже $x \notin A \Delta B$; якщо ж $x \notin A$, то $x \in B \backslash A$ і $x \in A \Delta B$. В обох випадках отримали суперечність щодо умови $A \Delta B = A$. (б) $B = E$; (в) $A = B$; (г) $A = \overline{B}$; (д) $A = B$; (е) $A \cap B = \varnothing$. **11.** (а) Нехай $A \cap B \subseteq C$. Візьмемо $x \in A \Rightarrow ((x \in A)$ і $((x \in B)$ або $(x \notin B))) \Rightarrow (((x \in A)$ і $(x \in B))$ або $((x \in A)$ і $(x \notin B))) \Rightarrow ((x \in A \cap B)$ або $(x \in \overline{B})) \Rightarrow ((x \in C)$ або $(x \in \overline{B})) \Rightarrow x \in \overline{B} \cup C$. Навпаки, нехай $A \subseteq \overline{B} \cup C$. $x \in A \cap B \Rightarrow ((x \in A)$ і $(x \in B)) \Rightarrow ((x \in \overline{B} \cup C)$ і $(x \in B)) \Rightarrow$ $\Rightarrow$ $(((x \in \overline{B})$ або $(x \in C))$ і $(x \in B)) \Rightarrow$ $(((x \in \overline{B})$ і $(x \in B))$ або $((x \in C)$ і $(x \in B))) \Rightarrow ((x \in \varnothing)$ або $(x \in C)) \Rightarrow x \in \varnothing \cup C \Rightarrow x \in C$.

1.3. 1. (а) $\{(1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4)\}$; (б) $\{(2, 1), (2, 2), (3, 1), (3, 2), (4, 1), (4, 2)\}$; (в) $\{(2, 2), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4), (4, 2), (4, 3), (4, 4)\}$; (г) $\{(3, 1), (3, 2), (4, 1), (4, 2)\}$; (д) $\{(1, 2, 1), (1, 2, 2), (1, 3, 1), (1, 3, 2), (1, 4, 1), (1, 4, 2), (2, 2, 1), (2, 2, 2), (2, 3, 1), (2, 3, 2), (2, 4, 1), (2, 4, 2)\}$; (е) $\{(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (2, 4)\}$. **2.** (а) Будь-які множини A і B такі, що $A \neq B$. Рівність виконується, коли $A = B$. (б) Нерівність виконується для всіх множин, бо кортежі $((a, b), c)$ і $(a, (b, c))$ нерівні. **3.** (а) $(x, y) \in (A \times B) \cap (B \times A) \Leftrightarrow ((x, y) \in A \times B$ і $(x, y) \in B \times A) \Leftrightarrow$ $\Leftrightarrow ((x \in A$ і $y \in B)$ і $(x \in B$ і $y \in A)) \Leftrightarrow (x \in A$ і $x \in B$ і $y \in A$ і $y \in B) \Leftrightarrow (x \in A \cap B$ і $y \in A \cap B) \Leftrightarrow (x, y) \in (A \cap B) \times (A \cap B)$. (б) $(x, y) \in (A \times B) \cup (B \times A) \Leftrightarrow$ $\Leftrightarrow ((x, y) \in A \times B$ або $(x, y) \in B \times A) \Leftrightarrow ((x \in A$ й $y \in B)$ або $(x \in B$ і $y \in A)) \Rightarrow$ $\Rightarrow ((x \in A \cup B$ і $y \in A \cup B)$ або $(x \in A \cup B$ і $y \in A \cup B)) \Leftrightarrow (x, y) \in (A \cup B) \times (A \cup B)$. **4.** З аналізу доведення включення попередньої задачі випливає, що такою умовою є $A = B$. Тоді з умов $x \in A \cup B$ і $y \in A \cup B$ матимемо відповідно $x \in A$ і $x \in B$ та $y \in A$ і $y \in B$, тобто всі переходи в доведенні рівносильні. **5.** (а) $(\varnothing, \varnothing, \varnothing)$, $(\varnothing, \{1\}, \varnothing)$, $(\varnothing, \{2\}, \varnothing)$, $(\varnothing, \{1, 2\}, \varnothing)$, $(\{1\}, \varnothing, \varnothing)$, $(\{2\}, \varnothing, \varnothing)$, $(\{1, 2\}, \varnothing, \varnothing)$, $(\{1\}, \{2\}, \varnothing)$, $(\{1\}, \{1, 2\}, \{1\})$, $(\{1, 2\}, \{1\}, \{1\})$, $(\{2\}, \{1, 2\}, \{2\})$, $(\{1, 2\}, \{2\}, \{2\})$, $(\{1, 2\}, \{1, 2\}, \{1, 2\})$. **6.** (а) $x \in \mathrm{Pr}_1(A \times B) \Leftrightarrow$ $\Leftrightarrow \exists y \in B$: $(x, y) \in A \times B \Leftrightarrow (x \in A$ і $y \in B) \Leftrightarrow x \in A$.

1.4. 1. (а) $\mathrm{Pr}_1 C_1 = \{a, b, d\}$, $\mathrm{Pr}_2 C_1 = \{1, 3, 4, 5\}$; (б) $C_1 \cup C_2 =$ $= \{(a, 1), (a, 3), (a, 4), (a, 5), (b, 1), (b, 2),$

$(b, 3), (c, 1), (d, 2), (d, 3), (d, 4), (d, 5)\}$; (в) $C_1 \cap C_2 = \{(a, 5), (b, 3), (d, 3)\}$; (г) $C_1 \backslash C_2 = \{(a, 1), (a, 3), (b, 1), (d, 3), (d, 5)\}$; (д) $\overline{C}_1 = \{(a, 2), (a, 4), (b, 2), (b, 4), (b, 5), (c, 1), (c, 2), (c, 3), (c, 4), (c, 5), (d, 1), (d, 2)\}$; (е) $C_1^{-1} = \{(1, a), (3, a), (5, a), (1, b), (3, b), (3, d), (4, d), (5, d)\}$. **2.** (а) $C_1 \circ D_1 = \{(a, \alpha), (a, \gamma), (b, \beta), (b, \gamma), (c, \alpha), (c, \beta), (d, \alpha), (d, \gamma)\}$; (б) $C_1 \circ C_2^{-1} = \{(a, b), (b, c), (b, d), (c, c), (c, d), (c, b), (d, b)\}$; (в) $\{(b, 4), (b, 1), (b, 2), (b, 3), (c, 2), (c, 4), (d, 2), (d, 4), (d, 1), (d, 3)\}$. **3.** Графіком відповідності C_1 є зовнішня частина круга радіусом 1 із центром у початку координат. Графік C_2 — це внутрішня частина (разом із межовими лініями) квадрата з вершинами в точках $(0, 3)$, $(3, 0)$, $(0, -3)$, $(-3, 0)$. Графіком C_3 є зовнішні частини двох парабол: $y^2 + x = 0$ й $y^2 - x = 0$. Позначимо через K квадрат із вершинами в точках $(0, 2)$, $(2, 0)$, $(0, -2)$, $(-2, 0)$. Графіком відповідності C_4 є усі точки смуги $x \in [-0{,}5; 0{,}5]$, що лежать за межами K, а також усі точки всередині квадрата K, для яких виконується $x \leq -0{,}5$ або $x \geq 0{,}5$. **4.** $C^{(2k-1)} = C$ і $C^{(2k)} = \{(x, y) \mid |x| = |y|, |x| \leq 1\}$, $k \in N$. **5.** $bc + d = ad + b$. **6.** (а) $A \subseteq B$. Нехай $i_A \subseteq A \times B$. Тоді для довільного елемента $x \in A$ маємо: $(x, x) \in i_A$, $(x, x) \in A \times B$ і $x \in B$. Навпаки, якщо $A \subseteq B$, то $(x, x) \in i_A \Rightarrow x \in A$ і $x \in B \Rightarrow (x, x) \in A \times B$. (б) $A \cap B = \varnothing$; (в) $A = B$; (г) $A \cap B = \varnothing$; (д) $A = B$; (е) $|A| = |B| = 1$. **7.** (а) Нехай $x \in \mathrm{Pr}_1 D \Rightarrow \exists y$: $(x, y) \in D \Rightarrow (x, y) \in C_1 \circ C_2 \Rightarrow (\exists z$: $(x, z) \in C_1$ і $(z, y) \in C_2) \Rightarrow x \in \mathrm{Pr}_1 C_1$. Аналогічно можна довести друге включення. (б) Нехай $C_1 \circ C_2 = \varnothing$. Припустимо, що існує елемент $z \in \mathrm{Pr}_2 C_1 \cap \mathrm{Pr}_1 C_2 \Rightarrow (z \in \mathrm{Pr}_2 C_1$ і $z \in \mathrm{Pr}_1 C_2) \Rightarrow (\exists x$: $(x, z) \in C_1$ і $\exists y$: $(z, y) \in C_2) \Rightarrow (x, y) \in C_1 \circ C_2$. Останнє суперечить умові, отже $\mathrm{Pr}_2 C_1 \cap \mathrm{Pr}_1 C_2 = \varnothing$. Навпаки, нехай $\mathrm{Pr}_2 C_1 \cap \mathrm{Pr}_1 C_2 = \varnothing$. Знову застосуємо метод доведення від супротивного і припустимо, що існує елемент $(x, y) \in C_1 \circ C_2$. Тоді $(\exists z$: $(x, z) \in C_1$ і $(z, y) \in C_2) \Rightarrow (z \in \mathrm{Pr}_2 C_1$ і $z \in \mathrm{Pr}_1 C_2) \Rightarrow z \in \mathrm{Pr}_2 C_1 \cap \mathrm{Pr}_1 C_2$. Отримали суперечність щодо умови. Отже, множина $C_1 \circ C_2$ не може містити елементів і є порожньою. **8.** (а) C — всюди визначена відповідність; (б) C — функціональна: (в) C — ін'єктивна; (г) C — сюр'єктивна; (д) C — взаємно однозначна, або бієктивна. **9.** (а) C_3, C_4, C_5; (б) C_2, C_3, C_5; (в) C_3, C_5; (г) C_3, C_4, C_5; (д) C_3, C_5. **10.** Нехай f — відображення з A в B. Візьмемо довільний елемент $x \in A$; тоді $(x, x) \in i_A$. Зі всюди визначеності f випливає, що існує такий елемент $y \in B$, що $(x, y) \in f \Rightarrow (y, x) \in f^{-1} \Rightarrow (x, x) \in f \circ f^{-1}$. Для доведення другого включення розглянемо елемент $(x, y) \in f \circ f^{-1} \Rightarrow (\exists z$: $(x, z) \in f^{-1}$ і $(z, y) \in f) \Rightarrow ((z, x) \in f$ і $(z, y) \in f)$. Із функціональності відображення f випливає, що $x = y$, отже $(x, y) \in i_B$. Навпаки, нехай виконуються включення $i_A \subseteq f \circ f^{-1}$ й $f^{-1} \circ f \subseteq i_B$. Розглянемо довільний елемент $x \in A \Rightarrow (x, x) \in i_A \Rightarrow (x, x) \in f \circ f^{-1} \Rightarrow (\exists y$: $(x, y) \in f$ і $(y, x) \in f^{-1})$. З останнього твердження випливає всюди визначеність f. Припустімо, що для деякого $x \in A$ існують елементи $(x, y) \in f$ і $(x, z) \in f$. Звідси $((x, y) \in f$ і $(z, x) \in f^{-1}) \Rightarrow (z, y) \in f^{-1} \circ f \Rightarrow (z, y) \in i_B \Rightarrow z = y$. Отже, відповідність f — функціональна. **11.** Нехай f_1, f_2, $f_1 \cup f_2$ — відображення з A в B. Припустімо, що $f_1 \neq f_2$. Тоді існує елемент $(x, y) \in f_1$ і $(x, y) \notin f_2$ або існує $(x, z) \in f_2$ і $(x, z) \notin f_1$. Розглянемо перший із цих випадків (для другого міркування аналогічні). З умов $(x, y) \notin f_2$ і f_2 — відображення, випливає, що існує такий елемент $y' \neq y$, що $(x, y') \in f_2$. Тоді матимемо, що $(x, y) \in f_1 \cup f_2$, $(x, y') \in f_1 \cup f_2$ і $y \neq y'$. Цей висновок суперечить умові функціональності $f_1 \cup f_2$. Отже, $f_1 = f_2$. Якщо $f_1 = f_2$, то $f_1 \cup f_2 = f_1$ і обернене твердження випливає з умов задачі. **12.** (а) Довільному кортежу $(a_{i_1}, a_{i_2}, \ldots, a_{i_k}) \in A^k$ поставимо у відповідність відображення $f \in A^{N_k}$, означене такою умовою: для $j \in N_k$ покладемо $f(j) = a_{i_j}$, $a_{i_j} \in A$, $j = 1, 2, \ldots, k$. (б) Довільній множині $B \in \beta(A)$ поставимо у відповідність відображення $f \in \{0, 1\}^A$, означене так: для $x \in A$ вважаємо $f(x) = 1$, якщо $x \in B$, а не то $f(x) = 0$. **13.** Див. доведення задачі 1.4.10.

1.5. 1. (а) Бієкцією між A і A є, наприклад, тотожна відповідність i_A. (б) Якщо $A \sim B$, то існує бієкція f між A і B. Неважко довести, що f^{-1} буде бієкцією між B і A; (в) Якщо f — бієкція між A і B, а g — бієкція між B і C, то можна довести, що відповідність $f \circ g$ є бієкцією між A і C. **2.** (а) $(2n - 1, n)$, $n \in N$; (б) $(n, (-1)^n [n/2])$, $n \in N$; (в) (n, kn), $n \in N$. **3.** (а) Так. Шуканою бієкцією може бути, наприклад, i_A. (б) Ні. Наприклад, $N \sim Z$, однак $N \neq Z$. **4.** (а) Доведемо це твердження методом від супротивного. Нехай A — скінченна множина, B — її власна підмножина й існує бієкція f між A і B. З умов випливає, що є такий елемент $a \in A$, що $a \notin B$; позначимо його a_1. Розглянемо послідовність елементів a_1, a_2, a_3, …, де $a_2 = f(a_1)$, $a_3 = f(a_2)$, … . Неважко довести, що $a_i \in A$ і $a_i \neq a_j$ для $i \neq j$, $i, j \in N$. Отже, множина елементів цієї послідовності є нескінченною, що суперечить умові скінченності множини A. **5.** Візьмемо такий довільний елемент b, що $b \notin A_1 \cup A_2$. Покладемо $B_1 = A_1 \times \{b\}$ і $B_2 = A_2 \times \{b\}$. **6.** Якщо f — бієкція між A і B, а g — бієкція між C і D, то можна довести, що відповідність, яка елементу $(x, y) \in A \times C$ ставить у відповідність елемент $(f(x), g(y)) \in B \times D$, є шуканою бієкцією між $A \times C$ і $B \times D$.

1.6. 1. Припустивши супротивне, отримаємо суперечність, тому що об'єднання скінченних множин є скінченною множиною. **2.** Нехай A — нескінченна множина. Тоді за теоремою 1.2 з неї можна виділити зліченну підмножину $B = \{b_1, b_2, \ldots, b_n, \ldots\}$. Неважко встановити бієкцію між множиною $(A \setminus B) \cup B = A$ і множиною $(A \setminus B) \cup \cup \{b_2, b_3, \ldots, b_n, \ldots\} = A'$, яка є власною підмножиною множини A. Для доведення достатності можна використати доведення задачі 1.5.4(а). **3.** Нехай A — зліченна або скінченна множина; тоді A рівнопотужна N або N_k. Позначимо через f і g відповідні бієкції між N і A чи N_k і A. У першому випадку шукане відображення з N в A збігається з f. У другому бієкцію g між N_k і A доозначимо до відображення g' з N в A, поклавши для всіх елементів n, більших k, $g'(n) = a$, де $a \in A$. Навпаки, нехай A є множиною значень деякого відображення f з N в A. З властивості функціональності f випливає, що його область значень має потужність не більшу, ніж потужність області визначення f, бо кожен елемент області визначення при відображенні має не більше одного образа. Отже, множина A є скінченною або зліченною. **4.** Позначимо дану зліченну множину через A, тоді множина всіх скінченних послідовностей довжини n, які складаються з елементів множини A, є A^n, а множина B всіх таких послідовностей дорівнює $A \cup A^2 \cup A^3 \cup \ldots \cup A^n \cup \ldots$. Оскільки кожна з множин A^n зліченна (див. наслідок 1.4.4), то за теоремою 1.4 множина B також зліченна. **5.** Зрозуміло, що множина C_n усіх n-елементних підмножин зліченної множини A має не більшу потужність, ніж потужність множини A^n усіх послідовностей довжини n, які складаються з елементів множини A, тому C_n є зліченною множиною. Множина C всіх скінченних підмножин множини A дорівнює $C_0 \cup C_1 \cup C_2 \cup \ldots$, тобто є зліченною. **6.** Відомо, що будь-який відкритий інтервал на дійсній прямій містить принаймні одну точку, якій відповідає раціональне число. Даній множині відкритих інтервалів можна поставити у взаємно однозначну відповідність множину A, що складається з таких точок. Оскільки потужність множини A не більша від потужності множини Q всіх раціональних чисел, то множина A є скінченною або зліченною. **7.** Див. доведення попереднього твердження.

1.7. 1. (а) Розмістити квадрати так, щоб їх площини були паралельні, а вершини знаходились на ребрах чотирикутної піраміди. Прямі, що проходять через вершину піраміди і перетинають площини квадратів, задають бієктивну відповідність між множинами точок цих квадратів. (б) Розмістити квадрат паралельно площині. Побудувати правильну чотирикутну піраміду, основою якої є даний квадрат, а вершина знаходиться на площині. Довільній точці A квадрата поставимо у відповідність точку B площини за таким правилом: 1) із точки A опустимо перпендикуляр на площину; 2) із центра квадрата проведемо пряму, що проходить через точку перетину перпендикуляра з бічною гранню піраміди; 3) B — це точка перетину цієї прямої з площиною. (в) Таку взаємно однозначну відповідність можна встановити, наприклад, за допомогою відображення f: $R \to R^+$, де $f(x) = a^x$, $a > 0$, $a \neq 1$. (г) Розмістити пряму так, щоб вона дотикалася до кола в точці, діаметрально протилежній вилученій точці A. Множина всіх пар точок кола і даної прямої, які знаходяться на прямих, проведених через точку A, утворює шукану бієкцію. (д) Див. розв'язок попередньої задачі. **2.** Нехай A_1, A_2, ... — довільні континуальні множини. Позначимо через B_n континуальну множину точок відрізка $[n-1; n)$. Тоді множина $A_1 \cup A_2 \cup \ldots \cup A_k$ рівнопотужна множині $B_1 \cup B_2 \cup \ldots \cup B_k$, тобто множині точок відрізка $[0; k)$, а множина $A_1 \cup A_2 \cup \ldots \cup A_k \cup \ldots$ рівнопотужна $B_1 \cup B_2 \cup \ldots \cup B_k \cup \ldots$, тобто множині точок півпрямої $[0; \infty)$. Обидві ці множини континуальні. **3.** Кожній такій послідовності $k_1, k_2, \ldots, k_n, \ldots$ поставимо у відповідність двійковий дріб із інтервалу $[0; 1)$ за правилом: після коми, що відділяє цілу частину 0 від дробової, записуємо $k_1 - 1$ одиницю, за нею пишемо 0, після нього — $k_2 - 1$ одиницю, знову 0 і т. д. Установлена взаємно однозначна відповідність між множиною A послідовностей і континуальною множиною чисел з інтервалу $[0; 1)$ доводить континуальність множини A. **4.** Помилка полягає в тому, що до складу булеана $\beta(N)$ окрім усіх скінченних підмножин входять також усі (не враховані в "доведенні") нескінченні підмножини множини N. **5.** Так. Наприклад, множина всіх кіл із центром у початку координат і радіусом r, де r пробігає множину дійсних чисел R (або інтервал $(0; 1)$). **6.** Ні. Кожен круг містить точку, обидві координати якої раціональні. Отже, потужність будь-якої множини таких кругів не перевищує потужність множини Q^2, яка є зліченною. **7.** Ні. Див. розв'язання попередньої задачі. **8.** Так. Див. 1.7.5.

1.8. 1. Це твердження є переформулюванням теореми 1.8 (Кантора—Бернштейна). **2.** Множина точок будь-якого квадрата $(a; b) \times \times (a; b)$ рівнопотужна множині точок квадрата $(0; 1) \times (0; 1)$, а множина точок відрізка $(c; d)$ рівнопотужна множині точок інтервалу $(0; 1)$. Оскільки відношення рівнопотужності множин транзитивне (див. задачу 1.5.1(в)), то для доведення цього твердження достатньо довести, що рівнопотужними є множини точок квадрата $(0; 1) \times (0; 1)$ та інтервалу $(0; 1)$. Позначимо першу з цих множин через A, другу — через B та скористаємося теоремою Кантора—Бернштейна. Нерівність $|B| \leq |A|$ випливає з того, що для $A' = \{0{,}5\} \times (0; 1)$ маємо $B \sim A'$ і $A' \subseteq A$. Для обґрунтування протилежної нерівності розглянемо таку відповідність f між A і B: довільній точці $(0{,}a_1a_2\ldots;\ 0,\ b_1b_2\ldots)$ квадрата A поставимо у відповідність точку $0{,}a_1b_1a_2b_2\ldots$ відрізка B. Ця відповідність є всюди визначеною, функціональною і сюр'єктивною, однак вона неін'єктивна (!). Усунувши "зайві" точки (тобто точки, що порушують умову ін'єктивності для f) із відрізка B, отримаємо підмножину B', для якої f буде бієкцією між A і B'. Отже, $|A| \leq |B|$. За теоремою Кантора—Бернштейна матимемо $|A| = |B|$. **3.** Вище доведено такі співвідношення: $(0; 1) \sim R$, $(0; 1) \times (0; 1) \sim R^2$ і $(0; 1) \times (0; 1) \sim (0; 1)$. Використовуючи властивості відношення рівнопотужності (див. задачі 1.5.1 і 1.5.6), отримаємо $R \sim R^2$, $R^2 \sim R^3$, $R^3 \sim R^4$, ... Отже, $R^n \sim R^m$ для довільних $n, m \geq 1$. **4.** Доведення можна виконати за допомогою методу математичної індукції. Для обґрунтування бази індукції ($n = 2$) й індукційного кроку використати задачу 1.8.2. **5.** (а) Покласти $B' = A$. (б) З умов задачі випливає, що існують такі підмножини $B' \subseteq B$ та $C' \subseteq C$, що $A \sim B'$ і $B \sim C'$. Позначимо відповідні бієкції через f і g. Нехай $C'' = g(f(A))$. Тоді можна довести, що відповідність $f \circ g$ є шуканою бієкцією між підмножиною $C'' \subseteq C$ і множиною A. (в) Нехай f і g — бієкції між A_1 і A_2 та B_1 і B_2, $B_1' \subseteq B_1$ і $B_1' \sim A_1$. Останню бієкцію між B_1' та A_1 позначимо через h. Тоді відповідність $g^{-1} \circ h \circ f$ є бієкцією між підмножинами $B_2' \subseteq B_2$ ($B_2' = g(B_1')$) та множиною A_2. (г) Нехай f — відображення з A на B. Утворимо підмножину $A' \subseteq A$, обравши для кожного $b \in B$ з $f^{-1}(b)$ по одному елементу c. Відповідність $\{(b, c) \mid$ для всіх $b \in B\}$ є бієкцією між B й A'. **6.** Нехай $a \in A$. Тоді покладемо $B' = \{a\} \times A \subseteq A \times \times A$. Очевидно, що $A \sim B'$. **7.** Позначимо через f і g бієкції між A_1 і B_1'

та A_2 і B_2' відповідно ($B_1' \subseteq B_1$, $B_2' \subseteq B_2$). (а) Відповідність $f \cup (g \setminus (A_1 \times B_2'))$ (або відповідність $g \cup (f \setminus (A_2 \times B_1')))$ є бієкцією між $A_1 \cup A_2$ і деякою підмножиною B'' множини $B_1' \cup B_2'$, тобто підмножиною множини $B_1 \cup B_2$. (б) Відповідність h, яка елементу $(a_1, a_2) \in A_1 \times A_2$ ставить у відповідність елемент $(f(a_1), g(a_2)) \in B_1' \times B_2'$, є бієкцією між $A_1 \times A_2$ і $B_1' \times B_2' \subseteq B_1 \times B_2$. **8.** Маємо $|f(A)| \le |A|$ (задача 1.8.5(г)) і $|A| < |\beta(A)|$ (теорема 1.9). Отже, $|f(A)| < |\beta(A)|$, і рівність множин $f(A)$ та $\beta(A)$ неможлива. **9.** (а) Континуум. Ця множина рівнопотужна множині всіх зліченних послідовностей натуральних чисел (див. задачу 1.7.3). (б) Континуум. Кожному монотонному відображенню f типу $N \to N$ поставимо у взаємно однозначну відповідність двійковий дріб з інтевалу $[0; 1)$ за таким правилом: пишемо 0 (ціла частина дробу), ставимо кому, далі записуємо $f(1) - 1$ одиницю, пишемо 0, за ним $f(2) - f(1)$ одиниць, пишемо 0, за ним $f(3) - f(2)$ одиниць, знову 0 і т. д. (в) Континуум. Запишемо дану множину у вигляді $D = R_1 \times R_2 \times \ldots \times R_i \times \ldots$, де $R_i = R$ для всіх $i \in N$. Для доведення континуальності D використаємо рівнопотужність пар множин D і R^N (узагальнення задачі 1.4.12(а)), R і N^N (див. задачу 1.8.9(а)), $N \times N$ і N (наслідок 1.4.4), а також твердження про рівнопотужність множин $(A^B)^C$ і $A^{B \times C}$, F^A та G^A, A^F і A^G для $F \sim G$, у справедливості яких пропонуємо переконатися самостійно. Тоді матимемо: $D \sim R^N$, $R^N \sim (N^N)^N$, $(N^N)^N \sim N^{N \times N}$, $N \times N \sim N$, $N^{N \times N} \sim N^N$, $N^N \sim R$. Отже, $D \sim R$. (г) Континуум. Кожній функції f типу $N \to \{0, 1\}$ поставимо у відповідність дріб $d = 0,a_1a_2a_3\ldots$ за таким правилом: якщо натуральне число $k \in \mathrm{Pr}_1 f$, то покладемо $a_k = f(k)$, а в іншому разі $a_k = 2$. Тоді даній множині функцій відповідатиме множина дробів, які є записом у трійковій системі числення всіх дійсних чисел з інтервалу $[0; 1]$. Зауважимо, що ця відповідність є всюди визначеною, функціональною та сюр'єктиною, але не є ін'єктивною (наприклад, двом різним функціям відповідатиме те саме число $0,1000\ldots = 0,0222\ldots$). Однак множина образів з $[0; 1]$, для яких порушується ін'єктивність, є підмножиною множини раціональних чисел із $[0; 1]$, отже є зліченною (див. доведення теореми 1.7). (д) Континуум. Див., наприклад, попередній пункт. З іншого боку, дана множина рівнопотужна $\beta(N)$ (див. задачу 1.4.12(б) та теорему 1.7). (е) Континуум. Нехай $r_1, r_2, \ldots, r_n, \ldots$ — якась послідовність раціональних чисел. Кожному неперервному відображенню f типу $R \to R$ поставимо у відповідність послідовність дійсних чисел $f(r_1), f(r_2), \ldots, f(r_n), \ldots$.

Ця відповідність взаємно однозначна, оскільки якщо розглянемо відображення f і g такі, що $f \neq g$, то існує дійсне число x, для якого $f(x) \neq g(x)$. Розглянемо послідовність $q_1, q_2, \ldots, q_n, \ldots$ раціональних чисел, що збігається до x. Тоді рівність $f(q_k) = g(q_k)$ виконується не для всіх k, оскільки в іншому разі, з урахуванням неперервності відображень f і g матимемо $f(x) = \lim\limits_{k\to\infty} f(q_k) = \lim\limits_{k\to\infty} g(q_k) = g(x)$. Отже, множина всіх неперервних відображень типу $R \to R$ рівнопотужна множині Q^N, яка є континуальною, оскільки $Q^N \sim N^N$ і $N^N \sim R$ (див. пункт (в) даної задачі). (є) Континуум, оскільки будь-яке монотонне відображення типу $R \to R$ однозначно визначається послідовністю своїх значень на зліченній множині точок з R: у точках розриву (див. задачу 1.6.7) і в точках $r_1, r_2, \ldots, r_i, \ldots$, де $r_i \in Q$, $i = 1, 2, \ldots$ (див. попередній пункт). (ж) Континуум (див. пункт (в)). (з) Більшою ніж континуум, бо власною підмножиною даної множини є множина $\{0, 1\}^R$, що рівнопотужна $\beta(R)$ (див. задачу 1.4.12(б), а також теорему 1.9). **10.** Власною підмножиною даної множини є, зокрема, множина $\{0, 1\}^{(0; 1)}$, рівнопотужна $\beta((0; 1))$ (див. 1.4.12(б)).

1.9. 1. (а) Множина (x, y), де x є батьком батька y. (б) Множина (x, y), де x — донька доньки y. (в) Множина (x, y), де x — батько, в якого є донька, а y — або мати цієї доньки, або $y = x$. (г) Множина (x, y), де x — людина жіночої статі й або x та y мають спільного батька, або $y = x$. (д) Множина (x, y), де x — онука, а y — її дідусь. (е) Порожня множина. (є) Множина (x, y), де y — людина жіночої статі й або x та y мають спільного батька, або $x = y$. (ж) Порожня множина. **2.** (а) $\forall y \in \mathrm{Pr}_2 R \Leftrightarrow \exists x: (x, y) \in R \Leftrightarrow \exists x: (y, x) \in R^{-1} \Leftrightarrow y \in \mathrm{Pr}_1 R^{-1}$. (б) Див. (а). (в) Нехай $\mathrm{Pr}_1 R = M$. Тоді для довільного елемента $x \in M$ послідовно маємо: $(x, x) \in i_M$ і $x \in \mathrm{Pr}_1 R \Leftrightarrow \exists y: (x, y) \in R \Leftrightarrow (y, x) \in R^{-1} \Rightarrow (x, x) \in R \circ R^{-1}$. Доведемо обернене твердження. Розглянемо довільний елемент $x \in M$. Тоді $(x, x) \in i_M$, і з умови матимемо $(x, x) \in R \circ R^{-1}$. Звідси отримаємо, що існує $y \in M$ такий, що $(x, y) \in R$ і $(y, x) \in R^{-1}$. Отже, $x \in \mathrm{Pr}_1 R$, тому доведено включення $M \subseteq \mathrm{Pr}_1 R$. Обернене включення випливає з означення проекції відношення. (г) Див. (в). **3.** (а) $\forall (x, y) \in (R_1 \circ R_2)^{-1} \Leftrightarrow$ $\Leftrightarrow (y, x) \in R_1 \circ R_2 \Leftrightarrow \exists z: (y, z) \in R_1, (z, x) \in R_2 \Leftrightarrow \exists z: (x, z) \in R_2^{-1}, (z, y) \in R_1^{-1} \Leftrightarrow$ $\Leftrightarrow (x, y) \in R_2^{-1} \circ R_1^{-1}$. (б) Нехай $R_1 \circ R_2 = \varnothing$. Припустімо супротивне, а саме, що існує елемент $z \in \mathrm{Pr}_2 R_1 \cap \mathrm{Pr}_1 R_2$. Тоді $z \in \mathrm{Pr}_2 R_1$ і $z \in \mathrm{Pr}_1 R_2$, тобто існують такі елементи x і y, що $(x, z) \in R_1$ і $(z, y) \in R_2$. Отримаємо $(x, y) \in R_1 \circ R_2$, що суперечить умові. Отже, $\mathrm{Pr}_2 R_1 \cap \mathrm{Pr}_1 R_2 = \varnothing$. Навпаки, нехай

$\mathrm{Pr}_2R_1 \cap \mathrm{Pr}_1R_2 = \varnothing$. Знову застосуємо метод доведення від супротивного: припустімо, що існує елемент $(x, y) \in R_1 \circ R_2$. Тоді матимемо $\exists z: (x, z) \in R_1$, $(z, y) \in R_2$ $\Rightarrow$ $(z \in \mathrm{Pr}_2R_1$ і $z \in \mathrm{Pr}_1R_2)$ $\Rightarrow$ $z \in \mathrm{Pr}_2R_1 \cap \mathrm{Pr}_1R_2$. Отримано суперечність, отже $R_1 \circ R_2 = \varnothing$. **4.** Для жодних. Припустивши існування деякого такого відношення R, матимемо: якщо $(x, x) \in R$, то $(x, x) \in R^{-1}$ і $(x, x) \notin \overline{R}$. Якщо ж $(x, x) \notin R$, то знову прийдемо до суперечності, тому що отримаємо $(x, x) \notin R^{-1}$ і $(x, x) \in \overline{R}$. **5.** (а) R_1, R_2, R_4; (б) R_5; (в) R_1, R_3; (г) R_2, R_5; (д) R_1. **6.** Усі діагональні елементи матриці рефлексивного відношення дорівнюють одиниці, а антирефлексивного — нулю. Матриця симетричного відношення є симетричною. У матриці антисиметричного відношення будь-які недіагональні елементи a_{ij} та a_{ji} $(i \neq j)$ або обидва нульові, або тільки один з них дорівнює одиниці. Графік рефлексивного відношення цілком включає в себе діагональ $y = x$ (для всіх x, що належать основній множині M даного відношення), а графік антирефлексивного відношення не містить жодного з елементів цієї діагоналі. Графік симетричного відношення є симетричним відносно діагоналі $y = x$, а графік антисиметричного відношення не містить жодної симетричної відносно діагоналі пари елементів (окрім елементів самої діагоналі). Граф рефлексивного відношення має петлі в кожній вершині, а граф антирефлексивного не має жодної петлі. У графі симетричного відношення для кожної дуги, що йде з i-ї в j-ту вершину, є протилежна дуга, яка веде з j в i. У графі антисиметричного відношення немає пар протилежних дуг. Якщо у графі транзитивного відношення є дві дуги, що ведуть з i в j та з j в k, то є й дуга, яка веде з i в k. **7.** (а) Нехай $(x, y) \in Q \circ R_1$. Тоді існує z такий, що $(x, z) \in Q$ і $(z, y) \in R_1$. Враховуючи умову, матимемо $(x, z) \in Q$ і $(z, y) \in R_2$, отже $(x, y) \in Q \circ R_2$. (б) Див. попереднє доведення. (в) $(x, y) \in R_1^{-1}$ $\Leftrightarrow$ $(y, x) \in R_1$ $\Rightarrow$ $(y, x) \in R_2$ $\Leftrightarrow$ $\Leftrightarrow$ $(x, y) \in R_2^{-1}$. **8.** (а) Нехай R — рефлексивне відношення на множині M. Тоді для довільного елемента $x \in M$ матимемо $(x, x) \in i_M$ (за означенням i_M) та $(x, x) \in R$ (за умовою), тобто $i_M \subseteq R$. Навпаки, припустімо, що $i_M \subseteq R$. Тоді для довільного $x \in M$ матимемо $(x, x) \in i_M$. Звідси, за умовою отримаємо $(x, x) \in R$, тобто R — рефлексивне. (б), (в) Доведення цих тверджень можна виконати аналогічно доведенню пункту (а). Крім того, можна довести, що обидва ці співвідношення рівносильні між собою та рівносильні умові (а). (див. задачу 1.2.5(а)). **9.** $R_1 = \{(1, 2), (2, 3), (3, 1)\}$, $R_2 = \{(2, 1), (3, 2), (1, 3)\}$. **10.** (а) Припустімо, що R — симетричне відношення. Розглянемо довільний елемент $(x, y) \in R$, тоді $(y, x) \in R$ і $(x, y) \in R^{-1}$. Аналогічно, якщо $(x, y) \in R^{-1}$, то $(y, x) \in R$ і $(x, y) \in R$. (б), (в) Обидві ці умови рівносильні умові (а), бо для довільних відношень R_1 і R_2 з $R_1 \subseteq R_2$ випливає $R_1^{-1} \subseteq R_2^{-1}$ та, крім того, $(R^{-1})^{-1} = R$. **11.** Нехай для симетричних відношень R_1 і R_2 їх композиція $R_1 \circ R_2$ є симетричним відношенням. Розглянемо довільний елемент $(x, y) \in R_1 \circ R_2$. Маємо таку послідовність рівносильних тверджень: $(y, x) \in R_1 \circ R_2 \Leftrightarrow (\exists z: (y, z) \in R_1$ і $(z, x) \in R_2) \Leftrightarrow (\exists z: (z, y) \in R_1$ і $(x, z) \in R_2) \Leftrightarrow (x, y) \in R_2 \circ R_1$. Отже, доведено рівність $R_1 \circ R_2 = R_2 \circ R_1$. Навпаки, припустімо, що для симетричних відношень R_1 і R_2 виконується $R_1 \circ R_2 = R_2 \circ R_1$. Розглянемо довільний елемент $(x, y) \in R_1 \circ R_2$, тоді матимемо $(x, y) \in R_2 \circ R_1 \Leftrightarrow (\exists z: (x, z) \in R_2$ і $(z, y) \in R_1) \Leftrightarrow (\exists z: (z, x) \in R_2$ і $(y, z) \in R_1) \Leftrightarrow (y, x) \in R_1 \circ R_2$. Симетричність $R_1 \circ R_2$ доведено. **12.** Нехай R — антисиметричне відношення на множині M і $(x, y) \in R \cap R^{-1}$, тоді $(x, y) \in R$ і $(x, y) \in R^{-1}$, тобто $(x, y) \in R$ і $(y, x) \in R$. З умови антисиметричності R матимемо $x = y$. Отже, $(x, y) \in i_M$. Навпаки, припустімо, що виконується включення $R \cap R^{-1} \subseteq i_M$. Розглянемо елементи $(x, y) \in R$ і $(y, x) \in R$. Звідси матимемо: $((x, y) \in R$ та $(x, y) \in R^{-1}) \Rightarrow (x, y) \in R \cap R^{-1} \Rightarrow (x, y) \in i_M \Rightarrow x = y$. Отже, R — антисиметричне відношення. **13.** Припустимо, що для антисиметричних відношень R_1 і R_2 об'єднання $R_1 \cup R_2$ є також антисиметричним відношенням. Розглянемо довільний елемент $(x, y) \in R_1 \cap R_2^{-1}$, тоді послідовно матимемо $((x, y) \in R_1$ і $(x, y) \in R_2^{-1}) \Leftrightarrow$ $\Leftrightarrow ((x, y) \in R_1$ і $(y, x) \in R_2) \Rightarrow ((x, y) \in R_1 \cup R_2$ та $(y, x) \in R_1 \cup R_2) \Rightarrow$ $\Rightarrow x = y \Rightarrow (x, y) \in i_M$. Навпаки, нехай для антисиметричних відношень R_1 і R_2 виконується включення $R_1 \cap R_2^{-1} \subseteq i_M$. Доведемо антисиметричність об'єднання $R_1 \cup R_2$. Для цього розглянемо елементи $(x, y) \in R_1 \cup R_2$ і $(y, x) \in R_1 \cup R_2$. Отримаємо таку послідовність тверджень: $(((x, y) \in R_1$ або $(x, y) \in R_2)$ і $((y, x) \in R_1$ або $(y, x) \in R_2)) \Leftrightarrow$ $\Leftrightarrow$ $(((x, y) \in R_1$ і $(y, x) \in R_1)$, або $((x, y) \in R_1$ і $(y, x) \in R_2)$, або $((x, y) \in R_2$ і $(y, x) \in R_1)$, або $((x, y) \in R_2$ і $(y, x) \in R_2)) \Rightarrow ((x = y)$ або $((x, y) \in R_1$ і $(x, y) \in R_2^{-1})$, або $((y, x) \in R_2^{-1}$ і $(y, x) \in R_1)$, або $(x = y)) \Rightarrow ((x = y)$ або $((x, y) \in R_1 \cap R_2^{-1}$ або $(y, x) \in R_1 \cap R_2^{-1}$ або $(x = y)) \Rightarrow ((x = y)$ або $(x, y) \in i_M$ або $(y, x) \in i_M$ або $(x = y)) \Rightarrow x = y$. **14.** Припустімо, що R — транзитивне відношення і $(x, y) \in R \circ R$. Тоді існує z такий, що $(x, z) \in R$ і $(z, y) \in R$. Отже, $(x, y) \in R$ (за властивістю транзитивності R). Навпаки, нехай виконується включення $R \circ R \subseteq R$. Розглянемо елементи $(x, y) \in R$ і

$(y, z) \in R$. Тоді $(x, z) \in R \circ R$, і з умови отримаємо $(x, z) \in R$. Отже, R — транзитивне відношення. **15.** Нехай для транзитивних відношень R_1 і R_2 виконується рівність $R_1 \circ R_2 = R_2 \circ R_1$. Розглянемо елементи $(x, y) \in R_1 \circ R_2$ і $(y, z) \in R_1 \circ R_2$. Тоді $(x, z) \in (R_1 \circ R_2) \circ (R_1 \circ R_2)$. З асоціативності операції композиції, умови цієї задачі, результату попередньої, а також задачі 1.9.7 отримаємо: $(R_1 \circ R_2) \circ (R_1 \circ R_2) = = R_1 \circ (R_2 \circ R_1) \circ R_2 = R_1 \circ (R_1 \circ R_2) \circ R_2 = (R_1 \circ R_1) \circ (R_2 \circ R_2) \subseteq R_1 \circ R_2$. Отже, $(x, z) \in R_1 \circ R_2$. **16.** Якщо $\Pr_1 R \neq M$, то із симетричності та транзитивності відношення R не випливає його рефлексивність. Наприклад, відношення $R = \{(1, 3), (3, 1), (1, 1), (3, 3)\}$ симетричне та транзитивне, але не рефлексивне на множині $M = \{1, 2, 3\}$. **17.** Рефлексивність T випливає з того, що для всіх дійсних x виконується нерівність $|x - x| < 1$, отже $(x, x) \in T$. Симетричність T випливає з рівності $|x - y| = |y - x|$. **18.** $R_1 = i_M \cup \{(1, 2), (2, 1)\}$ і $R_2 = i_M \cup \{(2, 3), (3, 2)\}$, $R_1 \circ R_2 = R_1 \cup R_2 \cup \cup \{(1, 3)\}$. **19.** Доведення проведемо методом математичної індукції. Для $k = 1$ твердження справедливе, оскільки за означенням $R^{(1)} = R$. Припустимо, що для рефлексивного та транзитивного відношення R виконується рівність $R^{(n)} = R$ для якогось $n \geq 1$. Для доведення рівності $R^{(n+1)} = R$ розглянемо довільний елемент $(x, y) \in R^{(n+1)}$. Оскільки $R^{(n+1)} = R^{(n)} \circ R$, то існує елемент z такий, що $(x, z) \in R^{(n)}$ і $(z, y) \in R$. За припущенням індукції маємо $(x, z) \in R$ і $(z, y) \in R$, отже $(x, y) \in R$ (тому що R — транзитивне). Таким чином, доведено включення $R^{(n+1)} \subseteq R$. Для обгрунтування оберненого включення візьмемо довільний елемент $(x, y) \in R$. За припущенням індукції $(x, y) \in R^{(n)}$, а з рефлексивності R випливає, що існує елемент $(y, y) \in R$. Отже, $(x, y) \in R^{(n)} \circ R = R^{(n+1)}$. **20.** Якщо $(a, b) \in R^{+}$, то за означенням транзитивного замикання існує послідовність елементів $a_1, a_2, \ldots, a_k$ така, що $a_1 = a$, $a_k = b$ і $a_1 R a_2$, $a_2 R a_3$, …, $a_{k-1} R a_k$. Звідси доходимо висновку, що $a_1 R^{(k-1)} a_k$, тобто $a R^{(k-1)} b$ для деякого k $(k \geq 2)$. Отже, $(a, b) \in R^{(1)} \cup R^{(2)} \cup \ldots \cup R^{(k)} \cup \ldots$. Навпаки, нехай $(a, b) \in R^{(1)} \cup R^{(2)} \cup \ldots \cup R^{(k)} \cup \ldots$. Тоді $(a, b) \in R^{(k)}$ для якогось k, $k = 1, 2, \ldots$. З означення відношення $R^{(k)}$ і властивостей операції композиції випливає, що існує набір елементів $z_1, z_2, \ldots, z_{k+1}$, для яких виконуються співвідношення $z_1 = a$, $z_{k+1} = b$ та $z_i R z_{i+1}$ для $i = 1, 2, \ldots, k$. Отже, $(a, b) \in R^{+}$. **21.** Якщо R транзитивне відношення, то можна довести, що $R^{(k)} \subseteq R$ для всіх $k \in N$ (див. задачі 1.9.14 і 1.9.7). Оскільки $R^{(1)} = R$, то $R^{+} = R^{(1)} \cup R^{(2)} \cup R^{(3)} \cup \ldots = R$ (див. попередню задачу та задачу 1.2.5(а)). Навпаки, нехай $R^{+} = R$. Розглянемо елементи $(x, y) \in R$ і $(y, z) \in R$. Тоді $(x, z) \in R^{(2)}$, отже $(x, z) \in R^{+}$ і тому $(x, z) \in R$. **22.** Якщо $k = 0$, то $C^{(0)}$ — це одинична матриця порядку n. Для $k \geq 1$ слід довести (застосувавши метод математичної індукції за параметром k), що елемент $c_{ij}^{(k)}$ i-го рядка і j-го стовпчика матриці $C^{(k)}$ дорівнює 1 тоді і тільки тоді, коли відповідний елемент a_{ij} матриці C^k не дорівнює 0. Отже, задача побудови матриці $C^{(k)}$ зводиться до обчислення k-го степеня матриці C відношення R.

1.10. 1. (а) Рефлексивність R випливає з того, що $a + b = b + a$, симетричність — із того, що коли $a + d = b + c$, то $c + b = d + a$. Для обгрунтування транзитивності розглянемо пари $((a, b), (c, d)) \in R$ і $((c, d), (e, f)) \in R$. Тоді, додавши почленно рівності $a + d = b + c$ та $c + f = d + e$, отримаємо $a + f = b + e$, тобто $((a, b), (e, f)) \in R$. (б) Рефлексивність Q випливає з того, що $ab = ba$, а симетричність — із того, що коли $ad = bc$, то $cb = da$. Транзитивність отримаємо, помноживши почленно відповідні рівності. **2.** (а) $[(1, 1)] = \{(1, 1)\}$, $[(2, 2)] = \{(1, 4), (4, 1), (2, 2)\}$, $[(4, 3)] = \{(1, 12), (12, 1), (2, 6), (6, 2), (3, 4), (4, 3)\}$, $[(1, 23)] = \{(1, 23), (23, 1)\}$, $[(6, 8)] = \{(1, 48), (48, 1), (2, 24), (24, 2), (3, 16), (16, 3), (4, 12), (12, 4), (6, 8), (8, 6)\}$. **3.** (а) Рефлексивність R випливає з того, що $m/m = 1 = 2^0$ і $0 \in Z$, а симетричність — із того, що коли $m/n = 2^k$, то $n/m = 2^{-k}$. Транзитивність отримаємо, помноживши почленно відповідні рівності. (б) Два. $[1]_R = [2]_R = [4]_R$, оскільки $1R2$ та $2R4$. Клас $[3]_R$ відмінний від усіх інших. (в) Три. $[6]_R = [12]_R = [24]_R = [48]_R$, $[7]_R = [28]_R$. Третім класом є $[42]_R$. **4.** (а) Відношення R означається через рівність множин, яка є рефлексивною, симетричною і транзитивною. (б) Чотири класи: $\{\varnothing, \{3\}\}$, $\{\{1\}, \{1, 3\}\}$, $\{\{2\}, \{2, 3\}\}$, $\{\{1, 2\}, \{1, 2, 3\}\}$. (в) $\{\{2, 3, 4\}, \{1, 2, 3, 4\}, \{2, 3, 4, 5\}, \{1, 2, 3, 4, 5\}\}$. (г) 8 класів, кожен з яких містить 4 елементи. (д) 2^m класів, у кожному з яких 2^{n-m} елементів. **5.** Нехай R_i, $i \in I$ довільна сукупність відношень еквівалентності на множині M і $D = \bigcap_{i \in I} R_i$. Оскільки для будь-якого $a \in M$ пара $(a, a) \in R_i$ для всіх $i \in I$, то $(a, a) \in D$. Отже, D — рефлексивне відношення. Якщо $(a, b) \in D$, то $(a, b) \in R_i$ для всіх $i \in I$. Кожне R_i симетричне, тому $(b, a) \in R_i$ для всіх $i \in I$, тобто $(b, a) \in D$, тому D — симетричне. Аналогічно, розглядаючи пари $(a, b) \in D$ і $(b, c) \in D$, доводимо транзитивність D. **6.** Див., наприклад, відповідь до задачі 1.9.18. **7.** Нехай відношення R_1, R_2 і $R_1 \cup R_2$ є еквівалентності. Доведемо рівність $R_1 \cup R_2 = R_1 \circ R_2$. Розглянемо елемент $(a, b)\in R_1\cup R_2$. Тоді $(a, b)\in R_1$ або $(a, b)\in R_2$. Із рефлексивності R_1 і R_2 маємо $(b, b)\in R_2$ і $(a, a)\in R_1$, отже в обох ситуаціях отримаємо $(a, b)\in R_1 \circ R_2$. Для доведення оберненого включення розглянемо довільний елемент $(a, b)\in R_1 \circ R_2$. Тоді існує елемент c такий, що $(a, c)\in R_1$ і $(c, b)\in R_2$. Використавши властивості операції об'єднання, маємо $(a, c)\in R_1\cup R_2$ і $(c, b)\in R_1\cup R_2$. За умовою, відношення $R_1\cup R_2$ транзитивне, тому $(a, b)\in R_1\cup R_2$. Необхідність доведено. Припустимо, що для еквівалентностей R_1 і R_2 виконується рівність $R_1\cup R_2 = R_1 \circ R_2$. Рефлексивність і симетричність об'єднання $R_1\cup R_2$ неважко вивести з рефлексивності і симетричності R_1 і R_2. Для доведення транзитивності відношення $R_1\cup R_2$ розглянемо пари $(a, b)\in R_1\cup R_2$ і $(b, c)\in R_1\cup R_2$. Останні умови рівносильні такій сукупності співвідношень: або 1) $(a, b)\in R_1$ та $(b, c)\in R_1$; або 2) $(a, b)\in R_1$ і $(b, c)\in R_2$; або 3) $(a, b)\in R_2$ і $(b, c)\in R_1$; або 4) $(a, b)\in R_2$ та $(b, c)\in R_2$. Для 1) або 4) з транзитивності R_1 і R_2 отримаємо відповідно $(a, c)\in R_1$ або $(a, c)\in R_2$, звідки матимемо $(a, c)\in R_1\cup R_2$. У випадку 2) маємо $(a, c)\in R_1 \circ R_2$, тому $(a, c)\in R_1\cup R_2$ (за умовою). Нарешті, для 3) послідовно отримаємо: $(b, a)\in R_2$ і $(c, b)\in R_1$ (симетричність R_1 і R_2) $\Rightarrow (c, a)\in R_1 \circ R_2 \Rightarrow$ $\Rightarrow (c, a)\in R_1\cup R_2 \Rightarrow (a, c)\in R_1\cup R_2$ (оскільки $R_1\cup R_2$ — симетричне відношення). **8.** Нехай $(a, b)\in R \circ R$. Тоді існує c такий, що $(a, c)\in R$ і $(c, b)\in R$. Оскільки R транзитивне, то $(a, b)\in R$. З іншого боку, якщо $(a, b)\in R$, то з рефлексивності R маємо $(b, b)\in R$ і, отже, $(a, b)\in R \circ R$. **9.** Рефлексивність композиції легко виводиться з рефлексивності R_1 і R_2: для довільного елемента $a\in M$ $(a, a)\in R_1$ і $(a, a)\in R_2$, отже, $(a, a)\in R_1 \circ R_2$. Для доведення симетричності розглянемо елемент $(a, b)\in R_1 \circ R_2$. Тоді послідовно маємо: $(a, b)\in R_1\cup R_2 \Leftrightarrow ((a, b)\in R_1$ або $(a, b)\in R_2) \Rightarrow ((b, a)\in R_1$ або $(b, a)\in R_2) \Leftrightarrow (b, a)\in R_1\cup R_2 \Leftrightarrow$ $\Leftrightarrow (b, a)\in R_1 \circ R_2$. Для обґрунтування транзитивності $R_1 \circ R_2$ візьмемо пари $(a, b)\in R_1 \circ R_2$ та $(b, c)\in R_1 \circ R_2$. Тоді, за умовою, $(a, b)\in R_1\cup R_2$ і $(b, c)\in R_1\cup R_2$. Далі див. доведення достатності в задачі 1.10.7. **10.** Неважко довести, що $R_1\cup R_2 \subseteq R_1 \circ R_2$ (див. задачу 1.10.7). Припустимо, що деяка еквівалентність Q містить $R_1\cup R_2$, тобто виконується включення $R_1\cup R_2 \subseteq$ $\subseteq Q$. Тоді $R_1 \circ R_2 \subseteq (R_1\cup R_2) \circ (R_1\cup R_2) \subseteq Q \circ Q = Q$ (див. задачі 1.9.7 і 1.10.8). **11.** (а) $Q = i_M\cup \{(2, 4), (4, 2), (3, 1), (1, 3)\}$. (б) $Q = i_M\cup \{(2, 1), (1, 2), (2, 3), (3, 2), (1, 3), (3, 1)\}$. **12.** Нехай R — відношення еквівалентності на множині M. Тоді $(x, y)\in (R \circ R^{-1})\cup i_M \Leftrightarrow ((x, y)\in R \circ R^{-1}$ або

$(x, y)\in i_M) \Leftrightarrow ((\exists z$: $(x, z)\in R$ і $(z, y)\in R^{-1})$ або $(x = y)) \Rightarrow (((x, z)\in R$ і $(y, z)\in R)$ або $(x = y)) \Rightarrow (((x, z)\in R$ і $(z, y)\in R)$ або $(x = y)) \Rightarrow ((x, y)\in R$ або $(x = y)) \Rightarrow (x, y)\in R$ (оскільки відношення R — рефлексивне). Припустимо, що $(x, y)\in R$, матимемо: $((x, y)\in R$ і $(y, y)\in R) \Rightarrow ((x, y)\in R$ і $(y, y)\in R^{-1}) \Rightarrow (x, y)\in R \circ R^{-1} \Rightarrow (x, y)\in R \circ R^{-1}\cup i_M$. Навпаки, нехай виконується дана рівність, доведемо еквівалентність R. Для довільного $x\in M \Rightarrow (x, x)\in i_M \Rightarrow (x, x)\in (R \circ R^{-1})\cup i_M \Rightarrow (x, x)\in R$. Отже, R — рефлексивне. Розглянемо елемент $(x, y)\in R \Leftrightarrow (x, y)\in$ $\in (R \circ R^{-1}) \cup i_M \Leftrightarrow ((x, y)\in R \circ R^{-1}$ або $(x, y)\in i_M) \Leftrightarrow ((\exists z$: $(x, z)\in R$ і $(z, y)\in R^{-1})$ або $(x = y)) \Leftrightarrow ((\exists z$: $(z, x)\in R^{-1}$ і $(y, z)\in R)$ або $(x = y)) \Leftrightarrow$ $((y, x)\in R \circ R^{-1}$ або $(y, x)\in i_M) \Leftrightarrow (y, x)\in (R \circ R^{-1}) \cup i_M \Leftrightarrow (y, x)\in R$. Таким чином, R — симетричне відношення. Для доведення транзитивності R розглянемо елементи $(x, y)\in R$ та $(y, z)\in R$. Оскільки R — симетричне, то $(x, y)\in R$ і $(z, y)\in R$. Звідси $((x, y)\in R$ і $(y, z)\in R^{-1}) \Rightarrow (x, z)\in R \circ R^{-1} \Rightarrow$ $\Rightarrow (x, z)\in (R\circ R^{-1})\cup i_M \Leftrightarrow (x, z)\in R$, тобто R — транзитивне. **13.** Див. означення класу еквівалентності $[x]_R$ і доведення теореми 1.10. **14.** Еквівалентність R випливає з того, що R означається через відношення рівності. Покладемо $f_1([x]_R) = f(x)$. Неважко переконатися, що f_1 — бієкція між A/R і $f(A)$. Тоді для довільного $x\in A$ $(\varepsilon \circ f_1)(x) = f_1(\varepsilon(x)) =$ $= f_1([x]_R) = f(x)$, тобто $f = \varepsilon \circ f_1$. **15.** (а) i_M. (б) M^2. **16.** Оскільки відношення R рефлексивне, то число $k - n$ дорівнює кількості недіагональних елементів відношення R. Це число парне внаслідок симетричності R.

1.11. 1. R — не лінійний порядок, оскільки непорівнюваними є, наприклад, числа 7 і 4, –11 і 10, 1 і 2 тощо. **2.** Ні. Відношення R — рефлексивне і транзитивне, але не антисиметричне, тому що з рівності $A\Delta B = C\Delta D$ не випливають рівності $A = C$ та $B = D$. Наприклад, $M = \{1, 2, 3, 4\}$, $A = \{1, 2, 3\}$, $B = \{1, 2, 4\}$, $C = \{1, 3\}$, $D = \{1, 4\}$. **3.** Усі три необхідні властивості відношення $\le$ неважко вивести з відповідних властивостей відношень $\le_A$ і $\le_B$. **4.** Для спростування цього твердження слід узяти такі дві пари елементів: $a \le_A b$, $a, b\in A$ $(a \ne b)$ та $d \le_B c$, $c, d\in B$ $(c \ne d)$. Тоді елементи $(a, c), (b, d)\in A \times B$ будуть непорівнюваними за відношенням $\le$. **5.** (а) R — частковий порядок. (б) R не є відношенням лінійного порядку. Наприклад, елементи $(\varnothing, M)$ і $(M, \varnothing)$ непорівнювані. **6.** Нехай R — частковий порядок на множині M. Розглянемо елемент $(x, y)\in R\cap R^{-1} \Leftrightarrow ((x, y)\in R$ і $(x, y)\in R^{-1}) \Leftrightarrow ((x, y)\in R$ і $(y, x)\in R) \Rightarrow x = y \Leftrightarrow$ $(x, y)\in i_M$. Якщо $(x, x)\in i_M$, то внаслідок рефлексивності R $(x, x)\in R$, тому

$(x, x) \in R^{-1}$. Отже, $(x, x) \in R \cap R^{-1}$. Навпаки, нехай R — транзитивне відношення і $R \cap R^{-1} = i_M$. Тоді для довільного $x \in M$ $(x, x) \in i_M \Leftrightarrow$ $\Leftrightarrow (x, x) \in R \cap R^{-1} \Rightarrow (x, x) \in R$; отже, R — рефлексивне. Розглянемо елементи $(x, y) \in R$ і $(y, x) \in R$. Тоді $(x, y) \in R$ і $(x, y) \in R^{-1}$. Звідси $(x, y) \in R \cap R^{-1} \Leftrightarrow$ $\Leftrightarrow (x, y) \in i_M \Leftrightarrow x = y$, тобто R — антисиметричне. **7.** Припустимо, що R_1, R_2 і $R_1 \cup R_2$ є відношеннями часткового порядку на M. Розглянемо довільний елемент $(x, y) \in R_1 \circ R_2 \cup R_2 \circ R_1$. Тоді послідовно матимемо: $(x, y) \in R_1 \circ R_2$ або $(x, y) \in R_2 \circ R_1 \Leftrightarrow$ (($\exists z$: $(x, z) \in R_1$ і $(z, y) \in R_2$) або ($\exists w$: $(x, w) \in R_2$ і $(w, y) \in R_1$)) $\Rightarrow$ ((($x, z) \in R_1 \cup R_2$ і $(z, y) \in R_1 \cup R_2$) або ((($x, w) \in R_1 \cup R_2$ і $(w, y) \in R_1 \cup R_2$. З обох останніх тверджень внаслідок транзитивності $R_1 \cup R_2$ матимемо $(x, y) \in R_1 \cup R_2$. Нехай $(x, y) \in R_1 \cap R_2^{-1} \Leftrightarrow$ (($x, y) \in R_1$ і $(x, y) \in R_2^{-1}$) $\Leftrightarrow$ (($x, y) \in R_1$ і $(y, x) \in R_2$) $\Rightarrow$ (($x, y) \in R_1 \cup R_2$ і $(y, x) \in R_1 \cup R_2$) $\Rightarrow x = y \Leftrightarrow (x, y) \in i_M$. Якщо ж розглянемо $(x, x) \in i_M$, то з рефлексивності R_1 і R_2 отримаємо, що (($x, x) \in R_1$ і $(x, x) \in R_2$) $\Leftrightarrow$ $\Leftrightarrow$ (($x, x) \in R_1$ і $(x, x) \in R_2^{-1}$) $\Leftrightarrow (x, x) \in R_1 \cap R_2^{-1}$. Необхідність доведено. Нехай R_1 і R_2 — відношення часткового порядку, і виконуються співвідношення $R_1 \circ R_2 \cup R_2 \circ R_1 \subseteq R_1 \cup R_2$ і $R_1 \cap R_2^{-1} = i_M$. Доведемо, що $R_1 \cup R_2$ — частковий порядок на M. Для довільного $x \in M$ $(x, x) \in R_1$, отже $(x, x) \in R_1 \cup R_2$, тобто $R_1 \cup R_2$ — рефлексивне. Розглянемо пари $(x, y) \in R_1 \cup R_2$ і $(y, x) \in R_1 \cup R_2$. Отримаємо таку рівносильну сукупність співвідношень: або 1) $(x, y) \in R_1$ і $(y, x) \in R_1$; або 2) $(x, y) \in R_1$ і $(y, x) \in R_2$; або 3) $(x, y) \in R_2$ і $(y, x) \in R_1$; або 4) $(x, y) \in R_2$ і $(y, x) \in R_2$. Для 1) і 4) безпосередньо з антисиметричності R_1 і R_2 матимемо $x = y$. Для 2) або 3) неважко вивести, що $(x, y) \in R_1 \cap R_2^{-1}$ або $(y, x) \in R_1 \cap R_2^{-1}$ відповідно (див. задачу 1.9.13). Отже, матимемо $(x, y) \in i_M$ або $(y, x) \in i_M$, тобто $x = y$. Антисиметричність $R_1 \cup R_2$ доведено. Для доведення транзитивності $R_1 \cup R_2$ розглянемо пари $(x, y) \in R_1 \cup R_2$ і $(y, z) \in R_1 \cup R_2$. У відповідній рівносильній сукупності співвідношень для випадків 1) або 4) з транзитивності R_1 або R_2 матимемо $(x, z) \in R_1$ або $(x, z) \in R_2$ відповідно, звідки $(x, z) \in R_1 \cup R_2$. У випадках 2) або 3) отримаємо $(x, z) \in R_1 \circ R_2$ або $(x, z) \in R_2 \circ R_1$ відповідно. Звідси $(x, z) \in R_1 \circ R_2 \cup R_2 \circ R_1$, тобто $(x, z) \in R_1 \cup R_2$. Твердження доведено. **8.** У множині M обов'язково існує принаймні один максимальний елемент a (див. задачу 1.11.10). Покладемо $f(a) = k$ і вилучимо елемент a з множини M (відповідно вилучимо з відношення часткового порядку, означеного на M, усі пари, що містять a). Множина $M \setminus \{a\}$ буде частково впорядкованою (довести це) і міститиме принаймні один максимальний елемент b. Означимо $f(b) = k - 1$ і вилучимо b з множини $M \setminus \{a\}$ і т. д. (Описана процедура називається *топологічним сортуванням* елементів

частково впорядкованої множини [24]). **9.** (а) Для будь-якого натурального n $(n, n+1) \in <$, однак $(n, n+1) \notin$ $\notin < \circ <$. (б) Нехай $(m, n) \in \le \circ < \Leftrightarrow$ ($\exists k$: $(m, k) \in \le$ і $(k, n) \in <$) $\Leftrightarrow$ ($m \le k$ і $k < n$) $\Rightarrow m < n \Leftrightarrow (m, n) \in <$. Розглянемо $(m, n) \in < \Rightarrow$ (($m, m) \in \le$ і $(m, n) \in <$) $\Rightarrow (m, n) \in \le \circ <$. (в) Включення $\le \circ \ge \subseteq N^2$ випливає з означення відношення на множині N. Нехай $(n, m) \in N^2$. Тоді виконується принаймні одна з нерівностей $n \le m$ або $n \ge m$. Якщо $n \le m$, тобто $(n, m) \in \le$, то $(m, m) \in \ge$, тому $(n, m) \in \le \circ \ge$. Якщо ж $n \ge m$, тобто $(n, m) \in \ge$, то $(n, n) \in \le$ і знову матимемо $(n, m) \in \le \circ \ge$. **10.** Для обгрунтування існування мінімального елемента у множині A використаємо метод доведення від супротивного. Припустимо, що A не містить жодного мінімального елемента. Тоді для довільного елемента $a \in A$ існуватиме $a_1 \in A$ такий, що $a_1 < a$. У свою чергу, для a_1 існуватиме $a_2 \in A$, для якого виконуватиметься $a_2 < a_1$, і т. д. Таким чином, множина A включатиме нескінченну підмножину $\{a_1, a_2, a_3, \ldots\}$, що суперечить умові. Аналогічно доводиться існування максимального елемента. **11.** Якщо частково впорядкована множина A має найменший елемент a, то він є єдиним мінімальним елементом у цій множині. Припустивши ж існування ще одного мінімального елемента b ($b \neq a$), отримаємо суперечність, бо тоді має виконуватись нерівність $a < b$. Нехай скінченна частково впорядкована множина A містить рівно один мінімальний елемент a. Доведемо, що a буде найменшим елементом множини A. Припустимо супротивне. Це означатиме, що у множині A існує елемент b, непорівнюваний з a. За умовою b не може бути мінімальним елементом, тому існує b_1 і $b_1 < b$. У свою чергу, для b_1 існуватиме b_2, для якого $b_2 < b_1$, і т. д. Зі скінченності множини A випливає, що ця послідовність має завершитись деяким b_k. Елемент b_k непорівнюваний з a, тому що $b_k < a$ суперечить умові мінімальності елемента a, а з $a < b_k$ можна вивести $a < b$, що суперечить непорівнюваності a і b. Таким чином, b_k є іншим мінімальним елементом множини A, що суперечить умові задачі. Отже, єдиний мінімальний елемент порівнюється з усіма елементами множини A, тому є її найменшим елементом. Для нескінченних частково впорядкованих множин дане твердження не є справедливим. Розглянемо, наприклад, множину $A = Z \cup \{0,5\}$, на якій означимо частковий порядок R так: на множині Z діє традиційне відношення порядку, а елемент 0,5 є порівнюваним лише сам із собою. Множина A, частково впорядкована за відношенням R, міститиме єдиний мінімальний (максимальний) елемент 0,5, однак не матиме ні найменшого, ні найбільшого елементів. **12.** (а), (б) Ні. (в) Так. **13.** Скінченні множини. Якщо M — нескінченна множина, на якій задано повний порядок R, то з множини M можна виділити зліченну підмножину $A = \{a_1, a_2, \ldots, a_n, \ldots\}$, для елементів якої виконуватимуться нерівності $a_1 < a_2 < \ldots < a_n < \ldots$. Тоді R^{-1} не буде повним порядком на M, оскільки для підмножини A не існуватиме найменшого елемента за відношенням R^{-1}. **14.** Припустимо, що таке монотонне відображення f існує. Позначимо через B підмножину всіх тих елементів a з A, для яких виконується $f(a) < a$. Тоді існує b, що є найменшим елементом множини B. Для елемента b маємо $f(b) < b$. Позначимо $c = f(b)$; тоді з $c < b$ й монотонності f отримаємо $f(c) < f(b)$, тобто $f(c) < c$. Отже, $c \in B$ й $c < b$. Це суперечить припущенню, що b — найменший елемент множини B. Зауважимо, що можна також довести, що для множини A не існує такого монотонного бієктивного відображення g множини A в себе, що для деякого $a \in A$ виконується $a < g(a)$. Отже, єдиним монотонним бієктивним відображенням цілком упорядкованої множини A в себе є тотожне відображення i_A.

1.12. 1. У діаграмі решітки для будь-якої пари вершин a та b виконуються такі умови: 1) існує принаймні одна вершина, з якої ведуть шляхи і в a, і в b; 2) множини вершин, що лежать на всіх шляхах, які ведуть в a або b, мають принаймні один спільний елемент (вважають, що з вершини a у вершину a веде шлях довжини 0, який складається з вершини a). Ці умови забезпечують існування точної верхньої грані для $\{a, b\}$. Аналогічно можна сформулювати умови, що відповідають ситуації існування $\inf\{a, b\}$. **2.** Припустимо, що в решітці L є максимальний елемент a, який не є найбільшим. Це означає, що в L існує елемент b, не порівнюваний з a. Нехай $c = \sup\{a, b\}$. Тоді $a \le c$ та $b \le c$. З умови максимальності a випливає $c = a$, тому $b \le a$. Остання нерівність суперечить непорівнюваності елементів a та b. **3.** Методом математичної індукції слід довести, що $\sup\{a_1, a_2, \ldots, a_n\} = \sup\{\ldots\{\sup\{\sup\{a_1, a_2\}, a_3\}, \ldots\}, a_n\}$ та $\inf\{a_1, a_2, \ldots, a_n\} = \inf\{\ldots\{\inf\{\inf\{a_1, a_2\}, a_3\}, \ldots\}, a_n\}$. Тоді $\sup\{a_1, a_2, \ldots, a_n\}$ є найбільшим, а $\inf\{a_1, a_2, \ldots, a_n\}$ — найменшим елементом решітки $L = \{a_1, a_2, \ldots, a_n\}$. **4.** (а), (б) випливають безпосередньо з означень. (в) з $a \le c$ і $b \le c$ випливає, що c є верхньою гранню множини $\{a, b\}$, отже $\sup\{a, b\} \le c$. (г) Див. (в). (д) Якщо $a \le b$, то b є верхньою гранню для $\{a, b\}$ і, очевидно, b є найменшою верхньою гранню. Якщо ж $\sup(a, b) = b$, то b — верхня грань для $\{a, b\}$, отже $a \le b$. (е) Див. (д). (є) Позначимо $f = \sup(c, d)$. З умов задачі та нерівностей $c \le f$ і $d \le f$ матимемо $a \le f$ і $b \le f$. Отже, f — верхня грань для $\{a, b\}$ і

$\sup(a, b) \le f$. (ж) Див. (є). (з) З $a \le c$ й $\inf(b, c) \le c$ отримаємо перше співвідношення $\sup(a, \inf(b, c)) \le c$ (див. (в)), а з $a \le a$ й $\inf(b, c) \le b$ — друге співвідношення $\sup(a, \inf(b, c)) \le \sup(a, b)$ (див. (є)). Застосувавши до обох цих співвідношень твердження з пункту (г), дістанемо $\sup(a, \inf(b, c)) \le \inf(\sup(a, b), c)$. (и) Нехай $\sup(a, b) = \inf(a, b)$. Позначимо цей елемент через c. Тоді з означень sup й inf маємо $a \le c$ й $b \le c$ та $c \le a$ й $c \le b$. Звідси $a = c$ та $b = c$, тобто $a = b$. Якщо $a = b$, то $\sup(a, a) = a$ й $\inf(a, a) = a$ (див. (а)). **5.** (а) Обидві рівності випливають безпосередньо з означень: елемент a, очевидно, є верхньою гранню для $\{a, a\}$ і він є точною верхньою гранню, бо для будь-якої іншої верхньої грані b множини $\{a, a\}$ виконується $a \le b$. (б) Для доведення першого співвідношення позначимо $\sup(a, \sup(b, c))$ через d, а $\sup(\sup(a, b), c)$ — через f. Тоді отримаємо: $a \le d$, $\sup(b, c) \le d \Rightarrow a \le d$, $b \le d$, $c \le d \Rightarrow \sup(a, b) \le d$, $c \le d \Rightarrow \sup(\sup(a, b), c) \le d \Rightarrow f \le d$. З іншого боку, знову використавши результати попередньої задачі (див. (б), (в)), маємо: $\sup(a, b) \le f$, $c \le f \Rightarrow a \le f$, $b \le f$, $c \le f \Rightarrow a \le f$, $\sup(b, c) \le f \Rightarrow \sup(a, \sup(b, c)) \le f \Rightarrow d \le f$. Отже, $d = f$. Друге співвідношення можна довести аналогічно. (в) Позначимо $\sup(\inf(a, b), a)$ через c. Тоді $\inf(a, b) \le c$ й $a \le c$. З іншого боку: $\inf(a, b) \le a$, $a \le a \Rightarrow \sup(\inf(a, b), a) \le a$ (див.1.12.4(в)) $\Rightarrow c \le a$. Отже, $c = a$. **6.** Якщо L — решітка, то існування $\sup A$ й $\inf A$ для довільної скінченної непорожньої множини A можна обґрунтувати, як у задачі 1.12.3. Обернене твердження випливає з того, що $\sup A$ й $\inf A$ існують, зокрема, і для двоелементних підмножин A. **7.** (а), (б), (в), (д) Так. (г) Ні. Наприклад, не існує точної верхньої грані для $\{R', R''\}$, де $R' = \{(a, b)\}$ й $R'' = \{(b, a)\}$ — антисиметричні відношення на множині $M = \{a, b\}$. **8.** (а), (б), (в) Ні. Для першої з цих множин не існує верхньої грані, для другої — нижньої грані. Третя множина не є повною решіткою (зокрема, тому, що включає в себе перші дві множини). **9.** Ні. Наприклад, для множини слів $\{a, aa, aaa, \ldots\}$, $a \in A$ не існує верхньої грані. **10.** Для довільної множини $A = \{d_1, d_2, \ldots, d_k\}$ дільників числа n існують числа $\sup A = \text{НСК}(d_1, d_2, \ldots, d_k)$ й $\inf A = \text{НСД}(d_1, d_2, \ldots, d_k)$, які є дільниками n. Нулем **0** цієї решітки буде число 1, а одиницею **1** — число n. **11.** Для обґрунтування цього твердження див. розв'язання задачі 1.12.3. **12.** (а) Використовуючи означення та результати задачі 1.12.4, маємо: 1) $\inf(a, b) \le a$, $\inf(a, c) \le a \Rightarrow \sup(\inf(a, b), \inf(a, c)) \le a$; 2) $\inf(a, b) \le b$, $\inf(a, c) \le c \Rightarrow \sup(\inf(a, b), \inf(a, c)) \le \sup(b, c)$. З 1) і 2) випливає (див.1.12.4(г)), що $\sup(\inf(a, b), \inf(a, c)) \le \inf(a, \sup(b, c))$. (б) Це твердження можна довести аналогічно попередньому.

2. Алгебричні системи

2.1. 1. (а), (в), (е) замкнені відносно $+$, (а) – (е) замкнені відносно $\times$. **2.** (а), (в), (е). **3.** (а) $0 \oplus k = k \oplus 0 = k$, $1 \oplus 1 = 2$, $1 \oplus 2 = 2 \oplus 1 = 0$, $2 \oplus 2 = 1$; $0 \otimes k = k \otimes 0 = 0$, $1 \otimes k = k \otimes 1 = k$, $2 \otimes 2 = 1$, $k = 0, 1, 2$. (б) $0 \oplus k = k \oplus 0 = k$, $1 \oplus 1 = 2$, $1 \oplus 2 = 2 \oplus 1 = 3$, $1 \oplus 3 = 3 \oplus 1 = 0$, $2 \oplus 2 = 0$, $2 \oplus 3 = 3 \oplus 2 = 1$, $3 \oplus 3 = 2$; $0 \otimes k = k \otimes 0 = 0$, $1 \otimes k = k \otimes 1 = k$, $2 \otimes 2 = 0$, $2 \otimes 3 = 3 \otimes 2 = 2$, $3 \otimes 3 = 1$, $k = 0, 1, 2, 3$. (г) $0 \oplus 0 = 1 \oplus 1 = 0$, $0 \oplus 1 = 1 \oplus 0 = 1$; $0 \otimes 0 = 0 \otimes 1 = 1 \otimes 0 = 0$, $1 \otimes 1 = 1$. **4.** (а), (б) Так. (в) $\otimes$ дистрибутивна відносно $\oplus$. **5.** (а) Так. (б) Ні, бо слова wp та pw, взагалі кажучи, різні для довільних $w, p \in A^*$. **6.** Результатом композиції $(\varphi \circ \psi) \circ \eta$ є підстановка $\alpha = \{(a, ((\varphi \circ \psi) \circ \eta)(a)) \mid a \in M\} = \{(a, \eta((\varphi \circ \psi)(a)) \mid a \in M\} = \{(a, \eta(\psi(\varphi(a))) \mid a \in M\}$, а результатом композиції $\varphi \circ (\psi \circ \eta)$ — підстановка $\beta = \{(a, (\varphi \circ (\psi \circ \eta))(a)) \mid a \in M\} = \{(a, (\psi \circ \eta)(\varphi(a))) \mid a \in M\} = \{(a, \eta(\psi(\varphi(a))) \mid a \in M\}$. Отже, $\alpha = \beta$. Для доведення некомутативності композиції розглянемо, наприклад, підстановки $\varphi = (1\ 3\ 2)$ та $\psi = (2\ 1\ 3)$ множини $M = \{1, 2, 3\}$. Тоді $\varphi \circ \psi = (2\ 3\ 1)$, а $\psi \circ \varphi = (3\ 1\ 2)$. **7.** $n!$ **8.** (а) n. (б), (в) Так. (г) Так. Якщо підстановці f відповідає зсув на k позицій, то, наприклад, для $m = \text{НСК}(k, n)/n$ матимемо $f^{(m)} = \varepsilon$. Зокрема, $f^{(m)} = \varepsilon$ для всіх $m = knt$, $t = 1, 2, \ldots$. **9.** Усюди визначепість і функціональність відповідності f випливають з означення. Для обгрунтування ін'єктивності припустимо, що для якихось $s, t \in M$ $(s > t)$ має місце рівність $f(s) = f(t)$. Це означає, що $ks = ml_1 + p$ та $kt = ml_2 + p$. Звідси $k(s - t) = m(l_1 - l_2)$. Оскільки k й m — взаємно прості числа, то $s - t$ має бути кратним m. Однак останнє неможливо, бо $s - t \le m - 1$. Неважко довести, що з ін'єктивності перетворення f скінченної множини M випливає, що f — сюр'єктивне, а отже, що f — підстановка. **10.** (а) Розглянемо послідовність підстановок $f^{(0)}, f^{(1)}, f^{(2)}, \ldots$. Зі скінченності множини P_M випливає, що існують числа k й l $(k < l)$, для яких $f^{(k)} = f^{(l)}$. Тоді $f^{(l-k)} = \varepsilon$. Більше того, $f^{(n(l-k))} = \varepsilon$ для всіх $n = 1, 2, \ldots$. (б) $h = f^{-1} \circ g$. (в) $w = g \circ f^{-1}$. **11.** Відношення R рефлексивне, бо для будь-якого $a \in M$ $f(a) = a$. Нехай aRb, тобто $f^{(m)}(a) = b$. Використовуючи результат попередньої задачі, оберемо число k, для якого $f^{(k)} = \varepsilon$ і $k \ge m$. Тоді $f^{(k-m)}(b) = a$, тобто bRa. Отже, відношення R — симетричне. Нехай aRb та bRc, тобто $f^{(m)}(a) = b$ і $f^{(k)}(b) = c$. Тоді $f^{(m+k)}(a) = c$, отже aRc і R — транзитивне.

2.2. 1. Покладемо $M = Z_{(k)}$, де $Z_{(k)}$ — множина всіх цілих чисел, кратних натуральному числу k. Відповідним ізоморфним відображенням є $\gamma(n) = kn$, $n \in Z$. **2.** (а) Покладемо $\gamma(x) = -x$, $x \in R^+$. Відображення γ взаємно однозначне, і для нього виконується співвідношення $\gamma(x + y) = -(x + y) = (-x) + (-y) = \gamma(x) + \gamma(y)$. (б) $\gamma(p) = 1/p$, $p \in M_1$. **3.** (а) Покладемо $\gamma\left(\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\right) = a + bi$. Неважко довести, що γ — бієкція, і для γ виконуються такі співвідношення: $\gamma(A + B) = \gamma(A) + \gamma(B)$, $\gamma(AB) = \gamma(A)\gamma(B)$, $A, B \in M_1$. (б) $\gamma\left(\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}\right) = a$ (в) $\gamma\left(\begin{pmatrix} a & b \\ 2b & a \end{pmatrix}\right) = a + b\sqrt{2}$. (г) $\gamma\left(\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\right) = (a, b)$. (д) $\gamma((a, b)) = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ **4.** (а) Ні, тому що не існує взаємно однозначного відображення для R і Q. (б) Ні. Припустимо, що існує ізоморфне відображення γ алгебри A_1 на алгебру A_2 та $\gamma(1) = a + bi$. Позначимо $\gamma(x) = c + di$, $x \in R$. Оскільки для будь-якого $x \in R$ мають виконуватися $\gamma(1 \times x) = \gamma(1) \times \gamma(x) = (a + bi)(c + di) = (ac - bd) + (ad + bc)i$ й $\gamma(1 \times x) = \gamma(x) = c + di$, то, поклавши, зокрема, $c = d = 1$, отримаємо $a = 1$ й $b = 0$. Отже, $\gamma(1) = 1 + 0i = 1$. Звідси $\gamma(2) = \gamma(1 + 1) = \gamma(1) + \gamma(1) = 1 + 1 = 2$. Позначимо $x = \gamma^{-1}(1 + i)$ й $y = \gamma^{-1}(1 - i)$. Тоді $\gamma(x + y) = \gamma(x) + \gamma(y) = 2$ та $\gamma(x \times y) = \gamma(x) \times \gamma(y) = 2$, тобто, з урахуванням бієктивності γ, $x + y = 2$ й $x \times y = 2$. Однак не існує дійсних чисел x та y, які задовольняють останні дві умови. (в) Ні. Якщо припустити існування ізоморфного відображення γ для A_1 і A_2, то $\gamma(0) = k$, $k \in N$. Тоді для довільного цілого числа n мають виконуватися $\gamma(n \times 0) = \gamma(n) \times \gamma(0) = \gamma(n) \times k$ та $\gamma(n \times 0) = \gamma(0) = k$, тобто $\gamma(n) = 1$ для всіх $n \in Z$, що неможливо для бієктивного відображення γ. (г) Ні. Застосувати метод, подібний до описаного в пункті (б). (д) Ні. Припустимо, що існує ізоморфне відображення γ алгебри A_1 на алгебру A_2 та $\gamma(0) = a$, $\gamma(1) = b$, $\gamma(-1) = c$. Тоді матимуть місце такі співвідношення: $\gamma(0 + 1) = \gamma(1) = \gamma(0) \times \gamma(1)$, $\gamma(0 + (-1)) = \gamma(-1) = \gamma(0) \times \gamma(-1)$, $\gamma(1 + (-1)) = \gamma(0) = \gamma(1) \times \gamma(-1)$, тобто $b = ab$, $c = ac$, $a = bc$. Однак не існує трьох попарно різних дійсних чисел, що задовольняють ці співвідношення. (е), (є) Ні (див. (д)). **5.** $\gamma(a + b\sqrt{2}) = a - b\sqrt{2}$, $a, b \in Q$. **6.** Якщо γ — ізоморфізм, то $\varphi = \gamma^{-1}$. Навпаки, нехай для гомоморфізму γ алгебри A_1 в A_2 існує гомоморфізм φ алгебри A_2 в A_1, для якого $\gamma \circ \varphi = i_{M_1}$ і $\varphi \circ \gamma = i_{M_2}$. Доведемо, що тоді відображення γ ін'єктивне та сюр'єктивне, тобто взаємно однозначне. Нехай $(x, y) \in \gamma$ та $(z, y) \in \gamma$. Оскільки за умовою для будь-якого $y \in M_2$ існує елемент w, для якого $(y, w) \in \varphi$, то $(x, w) \in \gamma \circ \varphi$ та $(z, w) \in \gamma \circ \varphi$. Отже, $(x, w) \in i_{M_1}$ і $(z, w) \in i_{M_1} \Rightarrow (x = w$ та $z = w) \Rightarrow x = z$, що доводить ін'єктивність γ. Сюр'єктивність γ випливає з такого ланцюга тверджень: $a \in M_2 \Rightarrow (a, a) \in i_{M_2} \Rightarrow (a, a) \in \varphi \circ \gamma \Rightarrow (\exists b \in M_1$: $(a, b) \in \varphi$ та $(b, a) \in \gamma)$. Отже, для довільного $a \in M_2$ існує елемент $b \in M_1$ такий, що $(b, a) \in \gamma$.

2.3. 1. Слід довести, що коли $m \equiv n \pmod p$ і $k \equiv l \pmod p$, то $m + k \equiv n + l \pmod p$ і $m \times k \equiv n \times l \pmod p$. **2.** Ізоморфним відображенням є $\gamma([k]_R) = k$, $k \in N_{(p)}$. **3.** Оскільки $x \; i_M \; y$ тоді й тільки тоді, коли $x = y$, то співвідношення $\varphi(a_1, a_2, \ldots, a_n) \, i_M \, \varphi(a_1, a_2, \ldots, a_n)$ матиме місце для будь-якої операції $\varphi \in \Omega$ та довільних елементів $a_1, a_2, \ldots, a_n \in M$. **4.** Відношення L рефлексивне, симетричне та транзитивне. Якщо $w_1 L w_2$ і $p_1 L p_2$, то $w_1 p_1 L w_2 p_2$, тому що з рівностей довжин слів w_1 і w_2 та p_1 і p_2 випливає рівність довжин слів $w_1 p_1$ і $w_2 p_2$. $A^*/L = \{A^0, A^1, A^2, \ldots\}$, а відповідну операцію на A^*/L означимо так: $A^n \circ A^m = A^{n+m}$, $n, m = 0, 1, 2, \ldots$. **5.** Див. розв'язання попередньої задачі. **6.** Відношення P є еквівалентністю. Окрім того, якщо для n-вимірних векторів v_1, v_2, w_1, w_2 виконується $v_1 \, P \, w_1$ і $v_2 \, P \, w_2$ (тобто перші координати векторів v_1 і w_1 та v_2 і w_2 збігаються), то вектори $v_1 + v_2$ і $w_1 + w_2$ та tv_1 і tw_1 також мають однакові перші координати, $t \in R$. Отже, $(v_1 + v_2) \, P \, (w_1 + w_2)$ і $tv_1 \, P \, tw_1$, тому P — конгруенція. Позначимо через W_a множину векторів з R^n, перша координата яких дорівнює a, $a \in R$. Тоді $R^n/P = \{W_a \mid a \in R\}$, а відповідні операції фактор-алгебри A/P означають так: $W_a + W_b = W_{a+b}$ та $tW_a = W_{ta}$, $a, b, t \in R$. **7.** Ізоморфним відображенням є $\gamma(W_a) = a$, $a \in R$. **8.** (а) Так. (б) Ні, бо з рівностей $\|a\| = \|b\|$ та $\|c\| = \|d\|$, взагалі кажучи, не випливає рівність $\|a + c\| = \|b + d\|$. **9.** Відношення R є еквівалентністю. Розглянемо елементи $(a, b) \, R \, (c, d)$ і $(e, f) \, R \, (g, h)$, тобто такі, що $ad = bc$ і $eh = fg$. Тоді $((a, b) \pm (e, f)) \, R \, ((c, d) \pm (g, h))$ або $(af \pm be, bf) \, R \, (ch \pm dg, dh)$, тому що $(af \pm be) \, dh = bf \, (ch \pm dg)$. Аналогічно $((a, b) \times (e, f)) \, R \, ((c, d) \times (g, h))$ або $(ae, bf) \, R \, (cg, dh)$, бо $aedh = bfcg$. Нарешті, якщо $e \neq 0$ та $g \neq 0$, то неважко обгрунтувати, що $((a, b) : (e, f)) \, R \, ((c, d) : (g, h))$. **10.** Ізоморфним відображенням є $\gamma((a, b)) = a/b$, $(a, b) \in M$.

2.4. 1. (б) Включення $[M_1] \subseteq [[M_1]]$ випливає безпосередньо з означення замикання множини. Якщо $a \in [[M_1]]$, то a є результатом деякої похідної операції f на аргументах множини $[M_1]$. Оскільки кожен із цих аргументів, у свою чергу, є результатом похідних операцій

на аргументах із M_1, то з означення суперпозиції випливає, що a є результатом операції з $[\Omega]_S$ на аргументах з M_1. Отже, $a \in [M_1]$. (в) Нехай $a \in [M_1]$. Це означає, що a є результатом деякої похідної операції f на аргументах із множини M_1. Оскільки $M_1 \subseteq M_2$, то всі ці аргументи є також елементами множини M_2. Отже, $a \in [M_2]$. (г) Якщо $a \in [M_1 \cap M_2]$, то a є результатом деякої похідної операції, аргументи якої належать множині $M_1 \cap M_2$, тобто кожен із цих аргументів є елементом і множини M_1, і множини M_2. Це означає, що $a \in [M_1]$ і $a \in [M_2]$, тому $a \in [M_1] \cap [M_2]$. Зауважимо, що обернене включення, взагалі кажучи, не виконується. (д) Нехай $a \in [M_1] \cup [M_2]$. Тоді $a \in [M_1]$ або $a \in [M_2]$. Якщо $a \in [M_1]$, то $a \in [M_1 \cup M_2]$. Аналогічно, якщо $a \in [M_2]$, то $a \in [M_1 \cup M_2]$. Отже, в обох випадках отримуємо, що $a \in [M_1 \cup M_2]$. (е) Якщо M_2 — замкнена множина, то $[M_2] = M_2$ і справедливість твердження випливає з (в). **2.** (а) $K_1 \subseteq K_2$ (див. 2.4.1(г)), але можливим є й строге включення $K_1 \subset K_2$. (б) $K_1 \not\subset K_2$. Наприклад, для алгебри $A = \langle N, \{\times\} \rangle$ покладемо $M_1 = \{3k \mid k \in N\} \cup \{3k + 2 \mid k \in N\}$ і $M_2 = \{6k + 2 \mid k \in N\} \cup \cup \{6k + 3 \mid k \in N\}$. Тоді, зокрема, $33 \in [M_1 \setminus M_2] = K_1$, однак $33 \notin [M_1] \setminus [M_2] = K_2$, оскільки $33 \in [M_1]$ і $33 \in [M_2]$. З іншого боку, $40 \in K_2$, однак $40 \notin K_1$. (в) $K_1 \subseteq K_2$, але можливо $K_1 \subset K_2$. Маємо $M_1 \cup (M_2 \cap M_3) = (M_1 \cup M_2) \cap (M_1 \cup M_3)$. Далі див. 2.4.1(г). (г) $K_2 \subseteq K_1$ або $K_2 \subset K_1$, тому що $M_1 \cap (M_2 \cup M_3) = (M_1 \cap M_2) \cup (M_1 \cap M_3)$. Далі див. 2.4.1(д). (д) $K_1 \not\subset K_2$. $M_1 \setminus (M_1 \cap M_2) = M_1 \setminus M_2$. Якщо $M_2 \subseteq M_1$, то $M_1 \cap M_2 = M_2$, далі див. пункт (б). **3.** $T = \{k + 1, k + 2, \ldots, 2k, 2k + 1\}$. **4.** Базисом алгебри A є, наприклад, $B = \{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 0, 1)\}$. **5.** Нехай P — скінченна система твірних алгебри A. Кожен з елементів $a \in P$ може бути породжений зі скінченної сукупності операндів, що належать системі твірних T (тобто a є результатом деякої похідної операції для аргументів із T). Множина всіх цих операндів утворює шукану скінченну систему твірних T'. **6.** Нехай T — скінченна система твірних алгебри A. Якщо T не є базисом, то існує елемент $b_1 \in T$, для якого $[T \setminus \{b_1\}] = M$. Якщо, у свою чергу, система твірних $T_1 = T \setminus \{b_1\}$ не є базисом, то вилучимо з неї відповідний елемент b_2. Продовжуючи цей процес, на певному кроці отримаємо скінченний базис T_k. **7.** Припустивши, що алгебра A скінченно-породжувана, отримаємо суперечність із результатом попередньої задачі.

2.5. 1. Теоретико-множинна операція перетину $\cap$ асоціативна та комутативна. Одиницею півгрупи A є множина M, тому що для будь-якої множини $B \in \beta(M)$ виконується $B \cap M = M \cap B = B$. Однак A не є півгрупою зі скороченням, бо з $B \cap C = D \cap C$, взагалі кажучи, не випливає рівність $B = D$. **2.** Операція композиції відношень на множині M всюди визначена й асоціативна, однак вона не комутативна. Одиницею півгрупи A є діагональне відношення i_M, тому що для будь-якого $R \in V$ виконується $R \circ i_M = i_M \circ R = R$. **3.** Операція $*$ всюди визначена, а в її асоціативності можна переконатися, розгортаючи вирази $(n * m) * k$ й $n * (m * k)$ за означенням і використовуючи властивості асоціативності й комутативності операцій додавання та множення для цілих чисел. З останніх властивостей випливає також комутативність операції $*$. Одиницею алгебри A є число 0. **4.** Одиницею алгебри A є число 1. Алгебра A не є півгрупою зі скороченням, тому що з $\max(k, n) = \max(m, n)$, взагалі кажучи, не випливає рівність $k = m$. **5.** Операція $*$ всюди визначена, а її асоціативність випливає з того, що порядок виконання операції $*$ не має значення, бо результатом виразу є a тоді й тільки тоді, коли всі його операнди дорівнюють a. (а) Так. (б) Так. Одиницею алгебри A є елемент b. (в) Ні, тому що $a * b = b * b$ та $b * a = b * b$, однак $a \neq b$. **6.** Ні. Наприклад, $(a * a) * b = a$, однак $a * (a * b) = b$. **7.** Для зручності не писатимемо знак операції $\times$. Із першого означення, зокрема, випливає, що для довільного елемента a існують такі елементи x та y, що $xa = e$ й $ay = e$. Із рівностей $xay = x(ay) = xe = x$ та $xay = (xa)y = ey = y$ отримаємо $x = y$. Навпаки, нехай виконуються умови другого означення. Тоді для довільних $a, b \in M$, елементами, що задовольняють умови $xa = b$ й $ay = b$, є $x = ba^{-1}$ й $y = a^{-1}b$. **8.** Нехай a' і a'' — елементи, обернені до a. Тоді $a''aa' = a''(aa') = a''e = a''$ й $a''aa' = (a''a)a' = ea' = a'$. Отже, $a'' = a'$. **9.** Якщо $ac = bc$, то $a = (bc)c^{-1} = b(cc^{-1}) = be = b$. Аналогічно можна довести властивість лівого скорочення. **10.** $(a \times b) \times (b^{-1} \times a^{-1}) = a \times (b \times b^{-1}) \times a^{-1} = a \times e \times a^{-1} = a \times a^{-1} = e$. Аналогічно, $(b^{-1} \times a^{-1}) \times (a \times b) = e$. Отже, $b^{-1} \times a^{-1}$ є оберненим елементом до $a \times b$. Із властивості єдиності оберненого елемента (див. 2.5.8) маємо $(a \times b)^{-1} = b^{-1} \times a^{-1}$. **11.** Припустімо супротивне: у мультиплікативній групі $A = \langle M, \{\times\}\rangle$ існують рівно два зазначені елементи a та b. Розглянемо елемент $c = a \times b \times a$. $c \neq e$, бо в іншому разі матимемо $a \times b = a^{-1} = a$, тобто $b = e$. Якщо припустити, що $c = a$, то $b \times a = e$, тобто $a = b^{-1} = b$. Нехай $c = b$. Тоді $a \times b = b \times a$. Позначимо $d = a \times b$. Зі зроблених припущень матимемо: $d \neq e$, $d \neq a$, $d \neq b$ та $d \times d = e$. Якщо ж $c \neq b$, то знову прийдемо до суперечності, бо $c \times c = e$. **12.** Для алгебри $A = \langle\{a, b\}, \{*\}\rangle$ операцію $*$ означимо такими співвідношеннями: $a * a = b * a = a$ й $a * b = b * b = b$. Операція $*$ всюди визначена й асоціативна, бо значення виразу $x * y * z$ в алгебрі A завжди дорівнює z для всіх $x, y, z \in \{a, b\}$ й тому не залежить від порядку виконання операцій у цьому виразі. Для $*$ виконуються умови лівого та не виконуються умови правого скорочення. Отже, A є скінченною півгрупою з лівим скороченням. Однак A не група, тому що ні елемент a, ні елемент b не є нейтральними, бо $a * b \neq b * a$. **13.** $A = \langle M, \{*\}\rangle$, де $e * x = x * e = x$, $x \in M$, $a * b = b * a = e$ й $a * a = b$, $b * b = a$. **14.** Операція $*$ всюди визначена й асоціативна. Одиницею алгебри A є елемент $(1, 0)$. Елементом, оберненим до (a, b), є $(c, -cb)$, де $c = 1/a$. Некомутативність групи A випливає з того, що рівність $ad + b = cb + d$ виконується не для всіх дійсних чисел a, b, c, d $(a \neq 0, c \neq 0)$. **15.** Нехай a, b — довільні елементи групи G, й $a \times b = c$. З умови випливає, що кожен елемент a обернений до себе. Тоді $a = c \times b$ та $b = a \times c$. Отже, $b \times a = (a \times c) \times (c \times b) = a \times (c \times c) \times b = a \times e \times b = a \times b$. **16.** (а) Операція звичайного числового множення має всі потрібні властивості. Одиницею групи A є 1, а кожен її елемент буде оберненим до себе. (б) Одиницею групи A є 1. Оберненим елементом до -1 є -1, до 1 — 1, до i — $-i$, до $-i$ — i. (в) Усі необхідні властивості операції додавання матриць ґрунтуються на відповідних властивостях операції додавання чисел. Нейтральним елементом (нулем) групи A буде нульова матриця, а оберненим елементом до будь-якої матриці T — матриця $-T$, тобто матриця, у якій усі елементи замінено на протилежні. (г) Одиницею групи A є одинична матриця, а елементом, оберненим до $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, — матриця $\begin{pmatrix} x & -y \\ y & x \end{pmatrix}$, де $x = a/(a^2 + b^2)$, $y = -b/(a^2 + b^2)$. **17.** Алгебра $\langle N \times N, \{+\}\rangle$ є комутативною півгрупою зі скороченням, бо всі необхідні властивості опера-ції $+$ випливають із відповідних властивостей операції додавання натуральних чисел. Так само можна обґрунтувати, що $\langle N \times N, \{\times\}\rangle$ — комутативна півгрупа. Дистрибутивність $\times$ відносно $+$ випливає з таких співвідношень: $(n, m) \times ((k, l) + (p, r)) = (n, m) \times (k + p, l + r) = (n(k + p) + m(l + r), n(l + r) + m(k + p)) = (nk + ml + np + mr, nl + mk + nr + mp) = (nk + ml, nl + mk) + (np + mr, nr + mp) = (n, m) \times (k, l) + (n, m) \times (p, r)$. **18.** Неважко довести, що $\langle N \times N, \{+\}\rangle$ — комутативна півгрупа зі скороченням, а $\langle N \times N, \{\times\}\rangle$ — комутативна півгрупа. Дистрибутивність операції $\times$ відносно $+$ випливає з відповідної дистрибутивності множення відносно додавання для натуральних чисел: $(n, m) \times ((k, l) + (p, r)) = (n, m) \times (k + p, l + r) = (n(k + p), m(l + r)) = (nk + np, ml + mr) = (nk, ml) + (np, mr) = (n, m) \times (k, l) + (n, m) \times (p, r)$. Одиницею A є елемент $(1, 1)$. **19.** (а) Із рівності $b + 0 = b$, яка виконується для довільного $b \in M$, маємо, що для будь-якого $a \in M$ $a \times (b + 0) = a \times b$, або (внаслідок дистрибутивності) $a \times b + a \times 0 = a \times b = a \times b + 0$. Користуючись властивістю скорочення для операції $+$, отримаємо $a \times 0 = 0$. Аналогічно можна довести, що $0 \times a = 0$. (б) $0 = 1 + (-1)$, тоді для довільного $a \in M$ маємо $0 = 0 \times a = (1 + (-1)) \times a = 1 \times a + (-1) \times a$. Отже, елемент $(-1) \times a$ є оберненим до a. Внаслідок єдиності оберненого елемента (див. 2.5.8) доходимо висновку, що $(-1) \times a = -a$. (в) Для довільних $a, b \in M$ маємо: $a + (-a) = 0 \Rightarrow (a + (-a)) \times b = a \times b + (-a) \times b = 0 \times b = 0$. Звідси доходимо висновку, що $(-a) \times b$ — елемент, обернений до $a \times b$, тобто $(-a) \times b = -(a \times b)$. Аналогічно доводимо, що $a \times (-b) = -(a \times b)$. Отже, $(-a) \times b = a \times (-b)$. (г) Користуючись результатом попереднього пункту, отримаємо $(-a) \times (-b) = a \times (-(-b)) = a \times b$, бо неважко переконатися, що будь-який елемент b збігається з елементом, оберненим до його оберненого, тобто з елементом $-(-b)$. **20.** (а) Так. (б) Ні. Не виконується умова дистрибутивності композиції $\circ$ відносно $+$, бо для довільних дійсних відображень f, g, h, узагалі кажучи, $(f + g) \circ h \neq f \circ h + g \circ h$. Це випливає з того, що рівність $h((f + g)(x)) = h(f(x) + g(x)) = h(f(x)) + h(g(x))$, яка має виконуватися для будь-якого $x \in R$, має місце лише для деяких відображень f, g, h. **21.** Ні, тому що в комутативній півгрупі $\langle \beta(M), \{\cup\} \rangle$ не виконуються умови скорочення: з $A \cup C = B \cup C$, взагалі кажучи, не випливає рівність $A = B$. **22.** $\langle M, \{ + \} \rangle$ — комутативна група, нулем якої є число $0 + 0\sqrt{2}$ або 0, а елементом, оберненим до $a + b\sqrt{2}$, — число $(-a) + (-b)$ $\sqrt{2}$ $\langle M \backslash \{0\}, \{ \times \} \rangle$ — комутативна група, одиницею якої є $1 + 0\sqrt{2}$ або 1, а елементом, оберненим до $a + b\sqrt{2}$, — число $(a/c) + (-b/c)\sqrt{2}$, де $c = a^2 - 2b^2$ (зауважимо, що не існує раціональних чисел a та b, для яких $a^2 - 2b^2 = 0$). Крім того, операція $\times$ дистрибутивна відносно $+$. **23.** (а) Ні. $\langle N_{(p)} \backslash \{0\}, \{\oplus\} \rangle$ не є групою, бо, наприклад, для елемента 2 не існує оберненого. (б) Так. $\langle N_{(p)}, \{\oplus\} \rangle$ — комутативна група, нулем якої є 0, елементом, оберненим до 0, є 0, а 1 і 2 є елементами, оберненими один до одного. $\langle N_{(p)} \backslash \{0\}, \{\otimes\} \rangle$ — комутативна група, одиницею якої є 1, $1^{-1} = 1$ та $2^{-1} = 2$. Операція $\otimes$ дистрибутивна відносно $\oplus$. (в) Так (див. розв'язання (б)). **24.** $\langle N_{(p)}, \{\oplus\} \rangle$ — комутативна група, нулем якої є 0, елементом, оберненим до 0, — 0, а елементом, оберненим до k ($k \neq 0$), — елемент $p - k$. $\langle N_{(p)}, \{\otimes\} \rangle$ — комутативна півгрупа з одиницею 1. Якщо p — складене число, то

$\langle N_{(p)} \backslash \{0\}, \{\otimes\} \rangle$ не є групою. Справді, нехай $p = km$. Тоді $k \otimes m = 0$. Отже, операція $\otimes$ є невизначеною для $k, m \in N_{(p)} \backslash \{0\}$. Операція $\otimes$ дистрибутивна відносно $\oplus$. Якщо ж p — просте число, то $\langle N(p) \backslash \{0\}, \{\otimes\} \rangle$ — комутативна група. Доведемо, що в цьому разі для будь-якого $k \in N_{(p)} \backslash \{0\}$ існує обернений елемент $m = k^{-1}$. Розглянемо сукупність елементів $1 \otimes k, 2 \otimes k, \ldots, (p - 1) \otimes k$. Жоден із цих елементів не дорівнює 0, бо жоден із відповідних добутків $k, 2k, \ldots, (p - 1)k$ не може бути кратним простому числу p. Жодні два $s \otimes k$ та $t \otimes k$ з цих елементів не збігаються між собою, бо в іншому разі за означенням операції $\otimes$ число $tk - sk = (t - s)k$ було б кратним p, що знову суперечило б простоті числа p. Отже, у зазначеній вище сукупності є всі елементи множини $\{1, 2, \ldots, p - 1\}$, зокрема, у цій сукупності є й шуканий елемент m, для якого $m \otimes k = 1$. Отже, $\langle N_{(p)}, \{\oplus, \otimes\} \rangle$ є полем, якщо p — просте число. **25.** $\langle R^2, \{ + \} \rangle$ — комутативна група, нулем якої є $(0, 0)$, а елементом, оберненим до (a, b), — елемент $(-a, -b)$. $\langle R^2 \backslash \{(0, 0)\}, \{ \times \} \rangle$ — комутативна група, одиницею якої є $(1, 0)$, а елементом, оберненим до (a, b), — елемент $(a/c, -b/c)$, де $c = a^2 + b^2$. Операція $\times$ дистрибутивна відносно $+$. **26.** Ізоморфне відображення $\gamma((a, b)) = a + bi$, $(a, b) \in R^2$. **27.** (а) Нехай $a \vee b = a \wedge b$. Тоді $a \vee (a \vee b) = (a \vee a) \vee b = a \vee b$ та $b \vee (a \vee b) = (b \vee a) \vee b = (a \vee b) \vee b = a \vee (b \vee b) = a \vee b$. Маємо $a \vee (a \vee b) = b \vee (a \vee b)$. З умов задачі та правил поглинання отримаємо $a \vee (a \vee b) = a \vee (a \wedge b) = a$ та $b \vee (a \vee b) = b \vee (a \wedge b) = b \vee (b \wedge a) = b$. Отже, $a = b$. Обернене твердження випливає з властивостей ідемпотентності операцій $\vee$ та $\wedge$. (б) Нехай $a \vee b = b$. Тоді $a \wedge (a \vee b) = a \wedge b$ і, за правилом поглинання, $a = a \wedge b$. Навпаки, якщо $a \wedge b = a$, то $b \vee (a \wedge b) = b \vee a$. Використовуючи комутативність обох операцій і друге правило поглинання, отримаємо $b = a \vee b$. **28.** Нехай $P = \langle M, \{\cup, \cap\}, \{ \leq \} \rangle$ — решітка, носієм якої є частково впорядкована множина M. Ототожнимо операції $\cup$ і $\vee$ та $\cap$ і $\wedge$. Доведемо, що для P виконуються всі властивості решітки, означеної в даному розділі. Усюди визначеність операцій $\cup$ та $\cap$ випливає з означення, а їх асоціативність і комутативність — із тверджень задачі 1.12.5(а, б). Ідемпотентність обох операцій обґрунтовано в задачі 1.12.4(а), а правила (або закони) поглинання — в 1.12.5(в). Навпаки, нехай задано решітку $R = \langle M, \{ \vee , \wedge \} \rangle$ як алгебру із зазначеними в цьому розділі властивостями. Означимо на множині M відношення $\leq$ таким чином: для $a, b \in M$ вважатимемо $a \leq b$ тоді й тільки тоді, коли $a \wedge b = a$. Доведемо, що $\leq$ — частковий порядок на M. Рефлексивність $\leq$ випливає з ідемпотентності операції $\wedge$. Нехай $a \le b$ та $b \le a$. Тоді $a \wedge b = a$ та $b \wedge a = b$. Оскільки операція $\wedge$ комутативна, то $a \wedge b = b \wedge a$ і $a = b$. Розглянемо співвідношення $a \le b$ та $b \le c$. Тоді $a \wedge b = a$ та $b \wedge c = b \Rightarrow a \wedge (b \wedge c) = a \Rightarrow (a \wedge b) \wedge c = a \Rightarrow a \wedge c = a \Leftrightarrow a \le c$. Доведемо, що в частково впорядкованій множині M для будь-яких $a, b \in M$ існують $\inf\{a, b\}$ та $\sup\{a, b\}$. Для цього доведемо, що $\inf\{a, b\} = a \wedge b$ та $\sup\{a, b\} = a \vee b$. Маємо: $(a \wedge b) \wedge a = a \wedge (b \wedge a) = a \wedge (a \wedge b) = (a \wedge a) \wedge b = a \wedge b$. Звідси випливає, що $a \wedge b \le a$. Аналогічно виведемо $a \wedge b \le b$. Отже, $a \wedge b$ — нижня грань для $\{a, b\}$. Нехай x — довільна нижня грань для $\{a, b\}$, тобто $x \le a$ й $x \le b$, що рівносильно $x \wedge a = x$ та $x \wedge b = x$. Тоді $x \wedge (a \wedge b) = (x \wedge a) \wedge b = x \wedge b = x$, тобто $x \le a \wedge b$. Таким чином, $a \wedge b$ — точна нижня грань $\{a, b\}$, або $\inf\{a, b\}$. Використовуючи результат попередньої задачі (див. 2.5.27(б)) та провівши аналогічні міркування стосовно $a \vee b$, доведемо, що $a \vee b$ — точна верхня грань для $\{a, b\}$. **29.** Означені операції $\vee$ та $\wedge$ всюди визначені, асоціативні та комутативні. Їх ідемпотентність випливає з рівностей НСК$(k, k) = k$ й НСД$(k, k) = k$. Доведемо перше правило поглинання, тобто рівність НСК$(k$, НСД$(k, m)) = k$. Якщо НСД$(k, m) = d$, то d — дільник числа k, тому НСК$(k, d) = k$. Аналогічно можна обгрунтувати й друге правило поглинання: НСД$(k$, НСК$(k, m)) = k$. Отже, A — решітка. Перший закон дистрибутивності в решітці A має такий вигляд: НСК$(k$, НСД$(l, m)) =$ = НСД(НСК(k, l), НСК$(k, m))$. Для доведення обох законів дистрибутивності слід використати таке твердження: якщо $t = p_1^{v_1} p_2^{v_2} \ldots p_r^{v_r}$ і $s = p_1^{w_1} p_2^{w_2} \ldots p_r^{w_r}$ — канонічні розклади натуральних чисел t й s на прості множники, то НСД$(t, s) = p_1^{z_1} p_2^{z_2} \ldots p_r^{z_r}$ і НСК$(t, s) = p_1^{u_1} p_2^{u_2} \ldots p_r^{u_r}$, де $z_i = \min(v_i, w_i)$ й $u_i = \max(v_i, w_i)$, $i = 1, 2, \ldots, r$. Нулем решітки A є дільник 1, а одиницею — число n.

2.6. 1. Для доведення слід переконатись, що для A справджуються всі шість пар аксіом булевої алгебри. **2.** Означимо на M відношення $\le$: вважаємо, що $a \le b$ тоді й тільки тоді, коли $a \wedge b = a$. Таким чином множина M перетворюється на решітку, де $\sup\{a, b\} = a \vee b$ й $\inf\{a, b\} = = a \wedge b$ (див. 2.5.28). Тоді неважко довести, що атом — це мінімальний елемент множини $M \backslash \{\mathbf{0}\}$. (а) Якщо x — мінімальний елемент, то $a = x$. В іншому разі існує $a_1 \le x$. Якщо a_1 — мінімальний елемент, то a_1 — шуканий атом. Інакше, існує $a_2 \le a_1$ і т. д. Внаслідок скінченності множини M цей процес завершується для деякого k, й a_k — шуканий атом. (б) Оскільки $a \wedge x = \inf\{a, x\}$, то $a \wedge x \le a$. З останньої нерівності й означення атома випливає, що або $a \wedge x = a$, або $a \wedge x = \mathbf{0}$. Крім того, обидва ці співвідношення не можуть виконуватися одночасно, бо $a \ne \mathbf{0}$. (в) Обидва ці співвідношення не можуть виконуватись одночасно, бо тоді отримаємо суперечність: $a \wedge a = = (a \wedge x) \wedge (a \wedge (\neg x)) = (a \wedge a) \wedge (x \wedge \neg x) = a \wedge \mathbf{0} = \mathbf{0}$. Якщо $a \le x$ (або $a \wedge x = a$), то твердження справедливе. Якщо ж не виконується $a \le x$, то $a \wedge x = \mathbf{0}$, бо $a \wedge x \le a$ (див. (б)). Тоді $a = a \wedge \mathbf{1} = a \wedge (x \vee (\neg x)) = = (a \wedge x) \vee (a \wedge (\neg x)) = \mathbf{0} \vee (a \wedge (\neg x)) = a \wedge (\neg x)$. Останнє співвідношення рівносильне нерівності $a \le \neg x$. **3.** Означимо на множині M частковий порядок $\le$ як у попередній задачі, перетворивши M на решітку. (а) $z \in \in \mathrm{At}(x \wedge y) \Leftrightarrow (z$ — атом і $z \le x \wedge y) \Leftrightarrow (z$ — атом і $(z \le x$ і $z \le y)) \Leftrightarrow (z \in \in \mathrm{At}(x)$ і $z \in \mathrm{At}(y)) \Leftrightarrow z \in \mathrm{At}(x) \cap \mathrm{At}(y)$. (б) Див. (а). (в) Зауважимо, що $\mathrm{At}(\mathbf{1})$ — множина всіх атомів алгебри A. $z \in \mathrm{At}(\neg x) \Leftrightarrow z \notin \mathrm{At}(x)$ (див. 2.6.2(в)) $\Leftrightarrow \Leftrightarrow z \in \mathrm{At}(\mathbf{1}) \backslash \mathrm{At}(x)$. (г) Нехай $\mathrm{At}(x) = \mathrm{At}(y)$ і припустимо, що $x \ne y$. Тоді принаймні одне з тверджень $x \le y$ або $y \le x$ хибне. Припустимо, що хибним є $x \le y$, тобто $x \wedge y \ne x$. Тоді $x \wedge (\neg y) \ne \mathbf{0}$, бо інакше для $x \wedge (\neg y) = \mathbf{0}$ отримаємо суперечність: $x = x \wedge \mathbf{1} = x \wedge (y \vee \neg y) = (x \wedge y) \vee (x \wedge \wedge (\neg y)) = (x \wedge y) \vee \mathbf{0} = x \wedge y$. Внаслідок 2.6.2(а) існує атом $a \le x \wedge (\neg y)$; звідси $a \in \mathrm{At}(x)$ і $a \in \mathrm{At}(\neg y)$, тобто $a \in \mathrm{At}(x)$ і $a \notin \mathrm{At}(y)$ (див. (а) і (в)). Отже, отримали нерівність $\mathrm{At}(x) \ne \mathrm{At}(y)$, яка суперечить припущенню. Аналогічні міркування проводимо для випадку, коли хибним є твердження $y \le x$. (д) Це твердження випливає безпосередньо з означень атома та множини $\mathrm{At}(x)$, бо $\mathrm{At}(x)$ складається з усіх таких атомів a, що $a \le x$. (е) Із того, що $a_i \le a_1 \vee a_2 \vee \ldots \vee a_k$, $i = 1, 2, \ldots, k$, і пояснень до попереднього пункту маємо $\{a_1, a_2, \ldots, a_k\} \subseteq \subseteq \mathrm{At}(a_1 \vee a_2 \vee \ldots \vee a_k)$. Для обгрунтування оберненого включення припустимо, що $a \in At(a_1 \vee a_2 \vee \ldots \vee a_k)$ й $a \ne a_i$, $i = 1, 2, \ldots, k$. Тоді $a \wedge a_i = \mathbf{0}$ для всіх $i = 1, 2, \ldots, k$ (див. 2.6.2(б)), отже $a \wedge (a_1 \vee a_2 \vee \ldots \vee a_k) = = (a \wedge a_1) \vee (a \wedge a_2) \vee \ldots \vee (a \wedge a_k) = \mathbf{0}$, що суперечить припущенню. **4.** Позначимо через L множину всіх атомів алгебри A. Тоді відображення $\gamma(x) = \mathrm{At}(x)$ є ізоморфізмом алгебри A в алгебру A'. Властивості ін'єктивності та сюр'єктивності γ випливають із тверджень (г) і (е) задачі 2.6.3. Отже, γ — бієкція. Співвідношення (2.1) з означення ізоморфізму виконуються згідно з твердженнями (а), (б), (в) попередньої задачі та рівностями $\gamma(\mathbf{0}) = \mathrm{At}(\mathbf{0}) = \varnothing$ і $\gamma(\mathbf{1}) = \mathrm{At}(\mathbf{1}) = L$. **5.** За доведеним у попередній задачі твердженням кожна з алгебр A_1 і A_2 ізоморфна деякій алгебрі множин $A_1' = \langle \beta(L_1), \{\cup, \cap, {}^{-}, \varnothing, L_1\}\rangle$ і $A_2' = \langle \beta(L_2), \{\cup, \cap, {}^{-}, \varnothing, L_2\}\rangle$ відповідно. З умови маємо $|\beta(L_1)| = |\beta(L_2)|$ або $|L_1| = |L_2|$. Будь-яка бієкція між множинами L_1 і L_2 є ізоморфізмом алгебр A_1' і A_2'. Із властивостей симетричності і транзитивності відношення ізоморфізму випливає ізоморфізм алгебр A_1 і A_2. **6.** Оскільки за теоремою Стоуна (задача 2.6.4) будь-яка скінченна булева алгебра $A = \langle M, \{ \vee , \wedge , \neg, \mathbf{0}, \mathbf{1}\}\rangle$ ізоморфна деякій булевій алгебрі множин $A' = \langle \beta(L), \{\cup, \cap, ^{-} , \varnothing, L\}\rangle$, то $|M| = |\beta(L)|$, а за теоремою 1.1 $|\beta(L)| = 2^{|L|}$. **7.** Наприклад, булеві алгебри множин $A = \langle \beta(L), \{\cup, \cap, ^{-} , \varnothing, L\}\rangle$ для будь-якої множини L, що складається з одного, двох або трьох елементів. **8.** $b = b \wedge \mathbf{1} = b \wedge (a \vee \neg a) = (b \wedge a) \vee (b \wedge \neg a) = \mathbf{0} \vee (b \wedge \neg a) = b \wedge \neg a$. У той же час $\neg a = \neg a \wedge \mathbf{1} = \neg a \wedge (a \vee b) = (\neg a \wedge a) \vee (\neg a \wedge b) = \mathbf{0} \vee (\neg a \wedge b) = \neg a \wedge b$. Отже, $b = \neg a$. **9.** (а) $\neg\neg a = \neg\neg a \wedge \mathbf{1} = \neg\neg a \wedge (a \vee \neg a) = (\neg\neg a \wedge a) \vee (\neg\neg a \wedge \neg a) = (\neg\neg a \wedge a) \vee \mathbf{0} = \neg\neg a \wedge a$. Крім того, $a = a \wedge \mathbf{1} = a \wedge (\neg\neg a \vee \neg a) = (a \wedge \neg\neg a) \vee (a \wedge \neg a) = (a \wedge \neg\neg a) \vee \mathbf{0} = a \wedge \neg\neg a$. Отже, $\neg\neg a = a$. (б) $\neg\mathbf{0} = \neg\mathbf{0} \vee \mathbf{0} = \mathbf{0} \vee \neg\mathbf{0} = \mathbf{1}$ (послідовно застосовано аксіоми 4, 2 та 6). $\neg\mathbf{1} = \neg\mathbf{1} \wedge \mathbf{1} = \mathbf{1} \wedge \neg\mathbf{1} = \mathbf{0}$. (в) Довести, що $\neg(\neg a \wedge \neg b) = a \vee b$, переконавшись у тому, що $(a \vee b) \vee (\neg a \wedge \neg b) = \mathbf{1}$ й $(a \vee b) \wedge (\neg a \wedge \neg b) = \mathbf{0}$. Відтак скористатися результатами задачі 2.6.8 і пункту (а). Друге співвідношення можна довести аналогічно. **10.** Рівносильність (а) і (б) обгрунтовано в задачі 2.5.27(б). Доведемо рівносильність (в) і (г). Нехай виконується (в). Тоді $\neg(a \wedge \neg b) = \neg\mathbf{0} = \mathbf{1}$, а $\neg(a \wedge \neg b) = \neg a \vee (\neg\neg b) = \neg a \vee b$ (див. 2.6.9(в, а)). Аналогічно можна обгрунтувати, що з (г) випливає (в). Відтак доведемо рівносильність сніввідношень (а) і (в). Нехай $a \wedge b = a$, тоді $a \wedge (\neg b) = (a \wedge b) \wedge (\neg b) = a \wedge (b \wedge \neg b) = a \wedge \mathbf{0} = \mathbf{0}$. Навпаки, нехай виконується (в). Тоді $a \wedge b = (a \wedge b) \vee \mathbf{0} = (a \wedge b) \vee (a \wedge (\neg b)) = a \wedge (b \vee \neg b) = a \wedge \mathbf{1} = a$. **11.** Нехай $(a \wedge (\neg b)) \vee (\neg a \wedge b) = \mathbf{0}$. Тоді $b = b \vee \mathbf{0} = b \vee ((a \wedge (\neg b)) \vee (\neg a \wedge b)) = (b \vee (a \wedge (\neg b))) \vee (\neg a \wedge b) = ((b \vee a) \wedge (b \vee \neg b)) \vee (\neg a \wedge b) = ((b \vee a) \wedge \mathbf{1}) \vee (\neg a \wedge b) = (b \vee a) \vee (\neg a \wedge b) = b \vee (a \vee (\neg a \wedge b)) = b \vee ((a \vee \neg a) \wedge (a \vee b)) = b \vee (\mathbf{1} \wedge (a \vee b)) = = b \vee (a \vee b) = a \vee b \vee b = a \vee b$. Аналогічно виведемо, що $a = a \vee \mathbf{0} = a \vee ((a \wedge (\neg b)) \vee (\neg a \wedge b)) = \ldots = a \vee b$. Отже, $a = b$. Навпаки, якщо $a = b$, то $(a \wedge (\neg b)) \vee (\neg a \wedge b) = (a \wedge \neg a) \vee (\neg a \wedge a) = \mathbf{0} \vee \mathbf{0} = \mathbf{0}$.

3. Математична логіка

3.1. 1. (а), (в), (д), (з), (и) — істинні, (б), (г) — хибні висловлення. **2.** (а), (в), (г), (д) — істинні, (б), (е), (є) — хибні висловлення. **3.** (а) Істинне висловлення $a \to b$, де a — висловлення "Трикутник є рівносторонній", b — "Трикутник є рівнобедрений". **4.** (а) 1. (б) 1.

5. (а) (00001011). (б) (01001000). **6, 7.** Побудувати відповідні таблиці істинності. **8.** (а), (б), (в), (е) — тавтології, (г), (д) — нейтральні. **9.** (а) Правильне твердження. (б) Неправильне твердження, бо воно не виконується для будь-якої нейтральної формули A. (в) Правильні обидва твердження. (г) Правильне твердження. Обернене твердження не виконується, наприклад, якщо A — тавтологія, а B — нейтральна формула або суперечність. (д) Правильне твердження. Обернене твердження не виконується, наприклад, якщо A — нейтральна формула або суперечність, а B — тавтологія. (е) Неправильне твердження, бо не виконується, наприклад, для двох суперечностей A і B. **10.** (а) Твердження не виконується для нейтральної формули A. (б) Правильне твердження. (в) Неправильне твердження, бо не виконується для будь-якої нейтральної формули A і $B = \neg A$. Обернене твердження правильне. (г) Правильне твердження. Обернене твердження не виконується, якщо A — нейтральна формула або тавтологія, а B — суперечність. (д) Твердження не виконується, наприклад, якщо A — суперечність, а B — довільна формула. (е) Див. 3.1.9(е). **11.** Ця формула набуває значення 1, якщо всі її операнди дорівнюють 1. **12.** Побудувати відповідні таблиці істинності. **13.** (а), (в), (г) — рівносильності. **14.** (в), (д), (е) — так. **15.** Припустімо, що формули A і B рівносильні. Тоді для будь-якого набору значень пропозиційних змінних формула $(A \sim B)$ матиме вигляд $0 \sim 0$ або $1 \sim 1$, тобто дорівнюватиме 1. Навпаки, нехай $(A \sim B)$ — тавтологія, однак формули A і B нерівносильні. Останнє означає, що існує набір значень, на якому формули A і B набувають різних значень. Тоді на цьому наборі формула $(A \sim B)$ дорівнюватиме значенню $0 \sim 1$ або $1 \sim 0$, тобто 0, що суперечить умові. **16.** (в), (д), (е) — так. **17.** A є суперечністю. **18.** Побудувати таблиці істинності формул A і B. **19.** A — суперечність. **20.** За означенням сильнішої формули A набуває значення 1 на всіх тих наборах, на яких B дорівнює 1. **21.** (а), (в), (д), (е) — так. **22.** Дана формула не є виконуваною тоді й тільки тоді, коли A_m — суперечність, а $A_1, A_2, \ldots, A_{m-1}$ — тавтології. Отже, множина формул $\{A_1, A_2, \ldots, A_m\}$ — суперечна. Обернене твердження не виконується, наприклад, якщо A_1 — суперечність, а решта формул — тавтології. **23.** Суперечна. **24.** *Б*. **25.** (а) *Б*. (б) *Б* сказав правду, а *А* і *В* — неправду. **26.** Несуперечна.

3.2. 1. (а), (б), (г) — так. **2.** (а) — так. **3.** (а) $F1$: $a \to b$; $F2$: $a \to c$; $F3$: $\mathrm{MP}(F1, A5) = (a \to c) \to (a \to (b \wedge c))$; $F4$: $\mathrm{MP}(F2, F3) = a \to (b \wedge c)$.

(б) $F1$: $S(a, b, c)(a, a, b)A5 = (a \to a) \to ((a \to b) \to (a \to (a \wedge b)))$; $F2$: $a \to a$ (див. приклад 3.2); $F3$: $\mathrm{MP}(F2, F1) = (a \to b) \to (a \to (a \wedge b))$; $F4$: $S(a, b)(b, a)A1 = b \to (a \to b)$; $F5$: b; $F6$: $\mathrm{MP}(F5, F4) = a \to b$; $F7$: $\mathrm{MP}(F6, F3) = a \to (a \wedge b)$; $F8$: a; $F9$: $\mathrm{MP}(F8, F7) = a \wedge b$. (в) $F1$: $a \to b$; $F2$: $a \to \neg b$; $F3$: $\mathrm{MP}(F1, A9) = (a \to \neg b) \to \neg a$; $F4$: $\mathrm{MP}(F2, F3) = \neg a$. (г) $F1$: $a \to b$; $F2$: $b \to c$; $F3$: $\mathrm{MP}(F1, A2) = (a \to (b \to c)) \to (a \to c)$; $F4$: $S(a, b)\ (b \to c, a)\ A1 = (b \to c) \to (a \to (b \to c))$; $F5$: $\mathrm{MP}(F2, F4) = a \to (b \to c)$; $F6$: $\mathrm{MP}(F5, F3) = a \to c$. **5.** (а) $F1$: $S(b)(a)A3 = a \wedge a \to a$. (б) $F1$: $S(b, c)(a, a)A5 = (a \to a) \to ((a \to a) \to (a \to (a \wedge a)))$; $F2$: $a \to a$; $F3$: $\mathrm{MP}(F2, F1) = (a \to a) \to (a \to (a \wedge a))$; $F4$: $\mathrm{MP}(F2, F3) = a \to (a \wedge a)$. (в) $F1$: $S(b, c)(a, a)A8 = (a \to a) \to ((a \to a) \to ((a \vee a) \to a))$; $F2$: $a \to a$; $F3$: $\mathrm{MP}(F2, F1) = (a \to a) \to ((a \vee a) \to a)$; $F4$: $\mathrm{MP}(F2, F3) = (a \vee a) \to a$. (г) $F1$: $S(b)(a)A6 = a \to (a \vee a)$. (д) $F1$: $S(b, c)(a, b)A5 = (a \to a) \to ((a \to b) \to (a \to (a \wedge b)))$; $F2$: $a \to a$; $F3$: $\mathrm{MP}(F2, F1) = (a \to b) \to (a \to (a \wedge b))$. (е) $F1$: $S(c)(a \vee c)A8 = (a \to (a \vee c)) \to ((b \to (a \vee c)) \to ((a \vee b) \to (a \vee c)))$; $F2$: $S(b)(c)A6 = a \to (a \vee c)$; $F3$: $\mathrm{MP}(F2, F1) = (b \to (a \vee c)) \to ((a \vee b) \to (a \vee c))$. (є) $F1$: $S(a)(c)A8 = (c \to c) \to ((b \to c) \to ((c \vee b) \to c))$; $F2$: $S(a)(c)(a \to a) = c \to c$; $F3$: $\mathrm{MP}(F2, F1) = (b \to c) \to ((c \vee b) \to c)$; $F4$: $S(b, c)(a, b)F3 = (a \to b) \to ((b \vee a) \to b)$. (ж) $F1$: $S(a, b)(a \to (a \vee b), a \wedge b)A1 = (a \to (a \vee b)) \to ((a \wedge b) \to (a \to (a \vee b)))$; $F2$: $\mathrm{MP}(A6, F1) = (a \wedge b) \to (a \to (a \vee b))$; $F3$: $S(a, b, c)(a \wedge b, a, a \vee b)A2 = ((a \wedge b) \to a) \to (((a \wedge b) \to (a \to (a \vee b))) \to ((a \wedge b) \to (a \vee b)))$; $F4$: $\mathrm{MP}(A3, F3) = ((a \wedge b) \to (a \to (a \vee b))) \to ((a \wedge b) \to (a \vee b))$; $F5$: $\mathrm{MP}(F2, F4) = (a \wedge b) \to (a \vee b)$. (з) $F1$: $S(c)(b \vee a)A8 = (a \to (b \vee a)) \to ((b \to (b \vee a)) \to ((a \vee b) \to (b \vee a)))$; $F2$: $S(a, b)(b, a)A7 = a \to (b \vee a)$; $F3$: $S(a, b)(b, a)A6 = b \to (b \vee a)$; $F4$: $\mathrm{MP}(F2, F1) = (b \to (b \vee a)) \to ((a \vee b) \to (b \vee a))$; $F5$: $\mathrm{MP}(F3, F4) = (a \vee b) \to (b \vee a)$. (и) $F1$: $S(a, c)$ $(a \wedge b, a)A5 = ((a \wedge b) \to b) \to (((a \wedge b) \to a) \to ((a \wedge b) \to (b \wedge a)))$. (і) $F1$: $S(b, c)(a \vee b, a)A5 = (a \to (a \vee b)) \to ((a \to a) \to (a \to ((a \vee b) \wedge a)))$; $F2$: $\mathrm{MP}(A6, F1) = (a \to a) \to (a \to ((a \vee b) \wedge a))$; $F3$: $a \to a$; $F4$: $\mathrm{MP}(F3, F2) = a \to ((a \vee b) \wedge a)$. (ї) $F1$: $S(b, c)(a \wedge b, a)A8$. **6.** (а) Обгрунтуємо вивідність: $(a \to b) \to a$, $a \to b \vdash b$. $F1$: $(a \to b) \to a$; $F2$: $a \to b$; $F3$: $\mathrm{MP}(F2, F1) = a$; $F4$: $\mathrm{MP}(F3, F2) = b$. (б) $a \to c$, $a \vdash b \vee c$. $F1$: $a \to c$; $F2$: a; $F3$: $\mathrm{MP}(F2, F1) = c$; $F4$: $S(a, b)(b, c)A7 = c \to (b \vee c)$;

$F5$: $MP(F3, F4) = b \vee c$. (в) Доведемо $\neg a \to b$, $\neg b \vdash a$. $F1$: $\neg a \to b$; $F2$: $\neg b$; $F3$: $S(a, b)(\neg b, \neg a)A1 = \neg b \to (\neg a \to \neg b)$; $F4$: $\mathrm{MP}(F2, F3) = \neg a \to \neg b$; $F5$: $S(a)(\neg a)A9 = (\neg a \to b) \to ((\neg a \to \neg b) \to \neg\neg a)$; $F6$: $\mathrm{MP}(F1, F5) = (\neg a \to \neg b) \to \neg\neg a$; $F7$: $\mathrm{MP}(F4, F6) = \neg\neg a$; $F8$: $\mathrm{MP}(F7, A10) = a$. (г) $a \to \neg b$, $b \vdash \neg a$. $F1$: $a \to \neg b$; $F2$: b; $F3$: $S(a, b)(b, a)A1 = b \to (a \to b)$; $F4$: $MP(F2, F3) = a \to b$; $F5$: $\mathrm{MP}(F4, A9) = (a \to \neg b) \to \neg a$; $F6$: $MP(F1, F5) = \neg a$. (д) $\neg b \to \neg a$, $a \vdash b$. $F1$: $\neg b \to \neg a$; $F2$: a; $F3$: $S(b)(\neg b)A1 = a \to (\neg b \to a)$; $F4$: $\mathrm{MP}(F2, F3) = \neg b \to a$; $F5$: $S(a, b)(\neg b, a)A9 = (\neg b \to a) \to ((\neg b \to \neg a) \to \neg\neg b)$; $F6$: $\mathrm{MP}(F4, F5) = (\neg b \to \neg a) \to \neg\neg b$; $F7$: $\mathrm{MP}(F1, F6) = \neg\neg b$; $F8$: $S(a)(b)A10 = \neg\neg b \to b$; $F9$: $\mathrm{MP}(F7, F8) = b$. (е) $a \vdash \neg\neg a$. $F1$: a; $F2$: $(a \to \neg b) \to (b \to \neg a)$; (див. (г)); $F3$: $S(a, b)(\neg a, a)F2 = (\neg a \to \neg a) \to (a \to \neg\neg a)$; $F4$: $a \to a$; $F5$: $S(a)(\neg a)F4 = \neg a \to \neg a$; $F6$: $\mathrm{MP}(F5, F3) = a \to \neg\neg a$; $F7$: $\mathrm{MP}(F1, F6) = \neg\neg a$. (є) $(a \to b) \to (a \to c)$, a, $b \vdash c$. $F1$: $(a \to b) \to (a \to c)$; $F2$: a; $F3$: b; $F4$: $S(a, b)(b, a)A1 = b \to (a \to b)$; $F5$: $\mathrm{MP}(F3, F4) = a \to b$; $F6$: $\mathrm{MP}(F5, F1) = a \to c$; $F7$: $\mathrm{MP}(F2, F6) = c$. (ж) $a \to (b \to c)$, $a \to b$, $a \vdash c$. $F1$: $a \to (b \to c)$; $F2$: $a \to b$; $F3$: a; $F4$: $\mathrm{MP}(F3, F1) = b \to c$; $F5$: $MP(F3, F2) = b$; $F6$: $\mathrm{MP}(F5, F4) = c$. **7.** Для обгрунтування (а) і (б) використати відповідно аксіоми $A8$ і $A5$. (в) Доведемо таку вивідність: Γ, $A \to B$, $\neg B \vdash \neg A$. $F1$: $A \to B$; $F2$: $\neg B$; $F3$: $S(a, b)(\neg B, A)A1 = \neg B \to (A \to \neg B)$; $F4$: $\mathrm{MP}(F2, F3) = A \to \neg B$; $F5$: $S(a, b)(A, B)A9 = (A \to B) \to ((A \to \neg B) \to \neg A)$; $F6$: $\mathrm{MP}(F1, F5) = (A \to \neg B) \to \neg A$; $F7$: $MP(F4, F6) = \neg A$. Отже, за МТД Γ, $A \to B \vdash \neg B \to \neg A$. **8.** (а) Див. 3.2.7(в). (б) $F1$: A; $F2$: $\neg A$; $F3$: $S(a, b)(\neg B, A)A9 = (\neg B \to A) \to ((\neg B \to \neg A) \to \neg\neg B)$; $F4$: $S(a, b)(A, \neg B)A1 = A \to (\neg B \to A)$; $F5$: $\mathrm{MP}(F1, F4) = \neg B \to A$; $F6$: $\mathrm{MP}(F5, F3) = (\neg B \to \neg A) \to \neg\neg B$; $F7$: $S(a, b)(\neg A, \neg B)A1 = \neg A \to (\neg B \to \neg A)$; $F8$: $MP(F2, F7) = \neg B \to \neg A$; $F9$: $\mathrm{MP}(F8, F6) = \neg\neg B$; $F10$: $S(a)(B)A10 = \neg\neg B \to B$; $F11$: $\mathrm{MP}(F9, F10) = B$. (в) $F1$: $A \wedge \neg A$; $F2$: $S(a, b)(A, \neg A)A3 = (A \wedge \neg A) \to A$; $F3$: $\mathrm{MP}(F1, F2) = A$; $F4$: $S(a, b)(A, \neg A)A4 = (A \wedge \neg A) \to \neg A$; $F5$: $\mathrm{MP}(F1, F4) = \neg A$. Далі див. (б). (г) 1) $A, \neg A \vdash B$ (див. (б)). 2) $B, \neg A \vdash B$ (див. 3.2.4(а, в)). З 1) і 2) матимемо $A \vee B$, $\neg A \vdash B$ (див. 3.2.7(а)). **9.** (а) 1) $a \wedge b, \neg a \vdash a$; 2) $a \wedge b, \neg a \vdash \neg a$; 3) $\neg a \vdash \neg(a \wedge b)$; 4) $a \wedge b, \neg b \vdash b$; 5) $a \wedge b, \neg b \vdash \neg b$; 6) $\neg b \vdash \neg(a \wedge b)$. Із 3) і 6) матимемо $\neg a \vee \neg b \vdash \neg(a \wedge b)$ (див. 3.2.7(а)). Отже, $\vdash (\neg a \vee \neg b) \to \neg(a \wedge b)$. (б) 1) $\neg(\neg a \vee \neg b), \neg a \vdash \neg a \vee \neg b$; 2) $\neg(\neg a \vee \neg b), \neg a \vdash \neg(\neg a \vee \neg b)$; 3) $\neg(\neg a \vee \neg b) \vdash \neg\neg a$; 4) $\neg\neg a \vdash a$;

5) $\neg(\neg a \vee \neg b) \vdash a$; 6) аналогічно отримаємо $\neg(\neg a \vee \neg b) \vdash b$; 7) із 5) і 6) матимемо $\neg(\neg a \vee \neg b) \vdash a \wedge b$ (див. 3.2.7(б)); 8) $\vdash \neg(\neg a \vee \neg b) \to$ $\to (a \wedge b)$; 9) $\neg(a \wedge b) \to \neg\neg(\neg a \vee \neg b)$ (див. 3.2.7(в)); 10) $\neg(a \wedge b) \vdash \neg\neg(\neg a \vee \neg b)$; 11) $\neg\neg(\neg a \vee \neg b) \vdash \neg a \vee \neg b$; 12) $\neg(a \wedge b) \vdash \neg a \vee \neg b$. Отже, $\vdash \neg(a \wedge b) \to (\neg a \vee \neg b)$. **10.** Модифікувавши доведення 3.2.8(в, б) (заміною $\neg B$ на b), можна довести теорему $(a \wedge \neg a) \to \neg b$. Тоді матимемо $F1$: $S(a)(a \wedge \neg a)A9 = ((a \wedge \neg a) \to b) \to$ $\to ((a \wedge \neg a) \to \neg b) \to \neg(a \wedge \neg a)$; $F2$: $(a \wedge \neg a) \to b$; $F3$: $(a \wedge \neg a) \to \neg b$; $F4$: MP($F2$, $F1$) $= ((a \wedge \neg a) \to \neg b) \to \neg(a \wedge \neg a)$; $F5$: MP($F3$, $F4$) $=$ $= \neg(a \wedge \neg a)$.

3.3. 1. (а) $P(k) = 1$ для $k = 6, 12, 18$. (б) $Q(k) = 1$ для $k = 2m$, $m = 1, 2, \ldots, 10$, і для $k = 3n$, $n = 1, 2, \ldots, 6$. **2.** Характеристичною множиною даного предиката є $\{(-2, -2), (-2, -1), (-2, 0), (-1, -2), (-1, -1),$ $(0, -2), (0, 2), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2)\}$. **3.** Позначимо через A, B, C множини істинності предикатів P, Q і R відповідно. (а) $\overline{A} \quad \cup B$. (б) $(\overline{A} \quad \cup B) \cap (A \cup \overline{B}) = (\overline{A} \quad \cap \overline{B}) \cup (A \cap B)$. (в) $\overline{A}$. (г) Збігається з (б). **4.** (а) Ч(a) $\wedge$ Ч(b) $\wedge$ (Б(c, a) $\wedge$ Б(c, b) $\vee$ М(d, a) $\wedge$ М(d, b)). (б) Ч(a) $\wedge$ Б(c, a) $\wedge$ М(d, a) $\wedge$ (Б(f, c) $\wedge$ Б(f, b) $\vee$ М(f, c) $\wedge$ М(f, b)). (в) Ч(a) $\wedge$ Б(a, c) $\wedge$ (Б(c, b) $\vee$ М(c, b)). (г) Ч(a) $\wedge \neg$Ч(b) $\wedge$ (Б(d, a) $\vee$ $\vee$ М(d, a) $\vee$ Б(f, b) $\wedge$ М(f, b)) $\wedge$ (Б(c, d) $\vee$ М(c, d) $\vee$ Б(c, f) $\vee$ М(c, f)).

3.4. 1. (а) $\exists x\ (\neg P(x))$. (б) $\neg(\exists x\ P(x))$. (в) $\neg(\forall x\ P(x))$. (г) $\forall x\ (\neg P(x))$. **2.** (а) x — вільна. (б) y — вільна змінна в $P(y)$. (в) y — вільна. (г) x — вільна змінна в $x > 3$. (д) Усі змінні зв'язані. **3.** (а) $\forall x\ \exists y\ (x < y)$. (б) $\forall z\ \exists y\ \forall x\ (xz = xy)$. (в), (г) Істинні висловлення. (д) $\forall q\ \exists p\ \forall x\ (x^2 +$ $+ px + q > 0)$. (е) Хибне висловлення. **4.** (а), (б) Хибні. **5.** Нехай M — множина прямих на площині. Тоді $\forall x\ \exists y\ (x \parallel y)$ — істинне висловлення, а $\exists y \forall x (x \parallel y)$ — хибне. Якщо Б(x, y) — предикат із 3.3.4, то $\forall y\ \exists x$ Б(x, y) — істинне висловлення, $\exists x\ \forall y$ Б(x, y) — хибне. **6.** Перше висловлення істинне, друге — хибне.

3.5. 1. Нехай M — довільна непорожня універсальна множина, $P(x)$ — довільний предикат на M. Якщо $P(x)$ тотожно істинний в області M, то формула $P(x) \vee \neg P(y)$ також буде тотожно істинною в M, в іншому разі існує $c \in M$, для якого $P(c) = 0$, або $\neg P(c) = 1$, тобто формула $P(x) \vee \neg P(c)$ істинна для всіх $x \in M$. Друга формула також є ЛЗЗ. **2.** (а) Застосуємо метод доведення від супротивного. Припустімо, що існує універсальна множина M, для якої дана формула хибна. Це означає, що $A = \exists x\ (P(x) \wedge Q(x))$ істинна, а $B = \exists x\ P(x) \wedge \exists x\ Q(x)$ хибна формула. Хибність B означає, що істинна принаймні одна з формул

$\forall x\ (\neg P(x))$ або $\forall(\neg Q(x))$. Однак істинність кожної з цих формул суперечить істинності A. Дане твердження можна довести й іншим способом. Нехай M — довільна непорожня предметна область, $P(x)$ і $Q(x)$ — будь-які предикати на M. Розглянемо випадок, коли антецедент $\exists x\ (P(x) \wedge Q(x))$ набуває значення 1. Тоді існує таке $c \in M$, що $P(c) \wedge Q(c) = 1$, тобто $P(c) = 1$ і $Q(c) = 1$. Отже, у цьому разі й консеквент $\exists x\ P(x) \wedge \exists x\ Q(x)$ істинний в M. Якщо антецедент хибний, ця формула істинна незалежно від значення консеквента. (б) Нехай M — довільна непорожня предметна область, $P(x)$ і $Q(x)$ — деякі предикати на M. Якщо $Q(x)$ виконуваний на M, то обидва операнди операції еквівалентності матимуть значення 1, отже дана формула істинна. Якщо ж предикат $Q(x)$ тотожно хибний в M, то значення операндів еквівалентності залежатимуть від предиката $P(x)$. Коли $P(x)$ тотожно істинний в M, то обидва ці операнди дорівнюватимуть 0, а не то (якщо предикат $P(x)$ виконуваний) — вони набувають значення 1. Отже, і за такої інтерпретації дана формула істинна. **3.** (б), (в) — тотожно істинні. Наприклад, розглянемо таку інтерпретацію формули (а): $M = \{a, b\}$ і для предикатів $P(x)$ і $Q(x)$ виконується $P(a) = 1$ і $P(b) = Q(a) = Q(b) = 0$. У цій інтерпретації дана формула буде хибною, бо $P(a) \wedge (P(b) \to Q(b)) \to$ $\to (Q(a) \vee Q(b)) = 0$. **4.** (а) Нехай M — довільна непорожня предметна область, $P(x)$ і $Q(x)$ — довільні предикати на M. Розглянемо ситуацію, коли формула A набуває значення 1. Це означає, що існує елемент $c \in M$, для якого $P(c) \vee Q(c) = 1$, тобто $P(c) = 1$ або $Q(c) = 1$. У цьому разі формула B також істинна. Припустимо, що формула A хибна в обраній інтерпретації. Тоді для всіх елементів x із M виконується $P(x) \vee Q(x) = 0$, тобто $P(x) = 0$ і $Q(x) = 0$. Звідси випливає, що й формула B набуває значення 0. Отже, доведено, що при довільній інтерпретації формули A і B набувають однакових значень, тобто вони рівносильні. **5.** (а), (б) $M = \{a, b\}$, $P(a) = Q(b) = 1$, $P(b) = Q(a) = 0$. (в) $M = \{a, b\}$, $P(a) = 1$, $P(b) = Q(a) = Q(b) = 0$. **6.** (а) A і B — рівносильні. (б) Контрприклад: $M = \{a\}$, $P(a) = Q(a) = 1$, $R(a) = 0$. (в) $M = \{a, b\}$, $P(a) = Q(a) = 0$, $P(b) = Q(b) = 1$. (г), (д) $M = \{a, b\}$, $P(a) = 1$, $P(b) = Q(a) = Q(b) = 0$. **7.** 1), 3), 4). Розглянути предикати $P(x)$: "x кратне 4" і $Q(x)$: "x — точний квадрат". Записати дані твердження у вигляді формул логіки предикатів. **8.** Введемо на множині M усіх чотирикутників такі предикати: $P(x)$ — "x — чотирикутник, вписаний у коло", $Q(x)$ — "x — прямокутник", $R(x)$ — "x

— паралелограм". Тоді припущення можна записати у вигляді таких предикатних формул: 1) $\exists x\ (P(x) \wedge Q(x))$, 2) $\forall x\ (Q(x) \to R(x))$, а висновок — у вигляді формули $\exists x\ (P(x) \wedge R(x))$. Із припущення 1) випливає, що для якогось $a \in M$ виконується $P(a) \wedge Q(a) = 1$, тобто $P(a) = 1$ та $Q(a) = 1$. Звідси і з припущення 2) маємо $R(a) = 1$. Отже, $P(a) \wedge R(a) = 1$, тому формула $\exists x\ (P(x) \wedge R(x))$ істинна в M. **9.** Означимо на множині N натуральних чисел такі предикати: $P(x)$ — "x кратне 51", $Q(x)$ — "x кратне 17", $R(x)$ — "x кратне 3", $T(x)$ — "сума цифр числа x кратна 3". Припущення матимуть такий вигляд: 1) $\forall x\ (P(x) \to (Q(x) \wedge R(x)))$, 2) $\forall x\ (R(x) \to T(x))$, 3) $\neg T(10712)$, а висновок — $\neg P(10712)$. Послідовно маємо: із 3) випливає $T(10712) = 0$, із 2) — $R(10712) = 0$, а з 1) — $P(10712) = 0$. Отже, $\neg P(10712)$ — істинне твердження. **10.** Означимо на сукупності множин такі предикати: $P(x)$ — "x — злічення", $Q(x)$ — "x — нескінченна", $R(x)$ — "x — незлічення", $T(x)$ — "x — скінченна". Припущення можна записати у вигляді: 1) $\forall x\ (P(x) \to Q(x))$, 2) $\forall x\ (R(x) \to Q(x))$, 3) $T(A)$, 4) $\forall x\ (T(x) \to \neg Q(x))$, а висновок — $\exists x\ (\neg P(x) \wedge \neg R(x))$. Тоді з 3) та 4) маємо $Q(A) = 0$, а з 1) і 2) — $P(A) = 0$ і $R(A) = 0$. Отже, $\neg P(A) \wedge \neg R(A) = 1$, тому формула $\exists x\ (\neg P(x) \wedge \neg R(x))$ істинна. **11.** Додамо до предикатів попередньої задачі предикат $S(x)$ — "множина x рівнопотужна деякій своїй власній підмножині". Припущення матимуть вигляд: 1) $\forall x\ (Q(x) \to S(x))$, 2) $\forall x\ (T(x) \to \neg S(x))$, а висновок: $\forall x\ (Q(x) \to \neg T(x)) \wedge \forall x\ (T(x) \to \neg Q(x))$. Розглянемо випадок, коли для деякої множини A $Q(A) = 1$. Тоді з 1) випливає, що $S(A) = 1$, а з 2) — $T(A) = 0$. Отже, для всіх таких множин A висновок буде істинним. Якщо ж $Q(A) = 0$, то обидві формули з висновку $Q(A) \to \neg T(A)$ і $T(A) \to \neg Q(A)$ істинні. **12.** Міркування неправильне. Означимо на множині студентів M такі предикати: $P(x)$ — "x може розв'язати дану задачу", $Q(x)$ — "здібний студент x здатен розв'язати цю задачу", $R(x)$ — "x — здібний студент". Припущення матимуть такий вигляд: 1) $\exists x\ P(x) \to (\exists x\ (R(x) \wedge Q(x)))$, 2) $R(a) \wedge (\neg P(a))$, де a — студент Петренко. Висновок можна записати у вигляді $\forall x\ (\neg P(x))$ або $\neg(\exists x\ P(x))$. Розглянемо ситуацію, коли в множині M існує інший здібний студент b, який може розв'язати цю задачу ($P(b) = Q(b) = R(b) = 1$). Тоді обидва припущення істинні, а висновок — ні. **13.** Означимо на сукупності множин такі предикати: $P(x)$ — "x — нескінченна", $Q(x)$ — "x — злічення", $R(x, y)$ — "потужність x не менша, ніж потужність y", $T(x, y)$ — "y — підмножина x". Посилки: 1) $\forall x\ (P(x) \to \exists y\ (Q(y) \wedge T(x, y)))$, 2) $P(I)$, де I — множина ірраціональних чисел, 3) $\forall x\ \forall y\ (T(x, y) \to R(x, y))$. Висновок: $\exists x\ (Q(x) \wedge R(I, x))$. З 1) і 2) випливає істинність формули $\exists y\ (Q(y) \wedge T(I, y))$, а з 3) — істинність $\forall y\ (T(I, y) \to R(I, y))$. Із першої з отриманих формул доходимо висновку, що існує така множина A, що $Q(A) = 1$ і $T(I, A) = 1$, а з другої — що $R(I, A) = 1$. Отже, $Q(A) \wedge R(I, A) = 1$, тому висновок істинний. **14.** Означимо такі предикати: $P(x)$ — "x — множина", $Q(x, y)$ — "потужність x більша, ніж потужність y", $R(x, y)$ — "y — підмножина x". Посилки: 1) $P(V)$, 2) $\forall x\ \exists y\ Q(y, x)$, 3) $\forall x\ \forall y\ (R(x, y) \to \to \neg Q(y, x))$, 4) $\forall x\ R(V, x)$. З 1), 3) та 4) випливає істинність формули $\forall y\ (\neg Q(y, V))$, а з 1) і 2) — істинність формули $\exists y\ Q(y, V)$. Однак одночасна істинність обох цих формул неможлива (перша з них рівносильна формулі $\neg(\exists y\ Q(y, V))$). **15.** Предикати на множині мешканців міста NN: $P(x)$ — "x — цирульник", $Q(x, y)$ — "x голить y". Посилка: $\forall x\ (P(x) \to \forall y\ (Q(x, y) \sim \neg Q(y, y)))$. Припустімо супротивне: $\exists x\ P(x)$, тобто для деякого a $P(a) = 1$. Тоді істинною буде формула $\forall y\ (Q(a, y) \sim \neg Q(y, y))$. Однак це неможливо, бо a є мешканцем міста NN, отже істинним має бути вираз $Q(a, a) \sim \neg Q(a, a)$. Отримана суперечність спростовує припущення про існування цирульника. **16.** Маємо таку послідовність рівносильних формул: $\exists x\,(\neg P(x)) \wedge \exists x\,(\neg Q(x))$, $\exists x\,(\neg P(x)) \wedge \exists y\,(\neg Q(y))$, $\exists x\,(\neg P(x) \wedge \exists y\,(\neg Q(y)))$, $\exists x\,(\exists y\,(\neg Q(y)) \wedge \neg P(x))$, $\exists x\,\exists y\,(\neg Q(y) \wedge \neg P(x))$, $\exists x\,\exists y\,(\neg(Q(y) \vee P(x)))$. Обидві останні формули є випередженими формами даної формули.

3.6. 1. Використати той факт, що для скінченної предметної області $M = \{a_1, a_2, \ldots, a_n\}$ предикатна формула $\forall x\ P(x)$ рівносильна формулі $P(a_1) \wedge P(a_2) \wedge \ldots \wedge P(a_n)$ для будь-якого предиката $P(x)$ на M. **2.** Нехай $M = \{a, b, c\}$ і $P(a) = P(c) = 1$, $P(b) = 0$. Тоді перша з формул рівносильна $P(y)$, а друга — $0 \to P(y)$. **3.** (а) Доведемо, що формула A є ЛЗЗ. Розглянемо довільну непорожню предметну область M і довільний предикат $P(x)$ на M. Якщо $P(x)$ тотожно істинний в M, то формула A тотожно істинна в M, в іншому разі (предикат $P(x)$ — не тотожно істинний в M), A рівносильна формулі $0 \to P(y)$, що також є тотожно істинною. Для доведення того, що формула B не є ЛЗЗ, можна розглянути, наприклад, інтерпретацію з попередньої задачі. (б) B — ЛЗЗ, A — ні. **4.** (а) Для довільної інтерпретації розглянути два випадки: $F(x)$ — тотожно істинна й $F(x)$ — не тотожно істинна формула. (б) Розглянути випадки, коли $F(x)$ — виконувана і коли $F(x)$ — тотожно хибна формула. **5.** (в), (г) — так. **6.** (а), (б) — ні. **7.** Застосуємо метод доведення від супротивного. Припустімо, що існує інтерпретація, у якій формула $A \to B(x)$ — ЛЗЗ, а формула $A \to \forall x\, B(x)$ — ні. Останнє означатиме, що існують таке значення c змінної x, для якого $B(c) = 0$, і такий набір значень d, на якому формула A набуває значення 1. Оскільки формула A не містить вільних входжень змінної x (тобто зміна значення x не впливає на значення формули A), то за вищезазначених умов формула $A \to B(x)$ також набуває значення 0, що суперечить припущенню про її тотожну істинність у даній інтерпретації. **8.** Див., наприклад, попереднє доведення. **9.** Для спростування твердження 7 розглянемо таку інтерпретацію: предметна область — множина N натуральних чисел, формула A — $x > 3$, формула B — $x > 2$ (отже, A містить вільне входження x). Тоді формула $A \to B(x)$ — тотожно істинна, а $A \to \forall x\, B(x)$ — не є тотожно істинною в N.

3.7. 1. (а) $\forall x\, ((21 \mid x) \to ((3 \mid x) \wedge (7 \mid x)))$. (б) $\forall x\, ((60 \mid x) \to ((10 \mid x) \wedge \wedge (6 \mid x))) \wedge \neg(\forall x\, (((10 \mid x) \wedge (6 \mid x)) \to (60 \mid x)))$. (в) Позначимо через $T(x)$ предикат "x — трансцендентне число", і через $I(x)$ — "x — ірраціональне число". $\forall x\, (T(x) \to I(x)) \wedge \neg(\forall x\, (I(x) \to T(x)))$, або $\forall x\, (T(x) \to \to I(x)) \wedge \exists x\, (I(x) \to \neg T(x))$. **2.** (а) $\forall x\, \forall y\, \exists z\, (x + y = z)$, $\forall x\, \forall y\, \exists z\, (x \times y = z)$. (б) $\neg(\forall x\, \forall y\, \exists z\, (x - y = z))$ або $\exists x\, \exists y\, \forall z\, (\neg(x - y = z))$. **3.** (а) $\forall x\, ((x \in A \Delta B) \sim \sim (((x \in A) \wedge \neg(x \in B)) \vee ((x \in B) \wedge \neg(x \in A))))$. (б) $\forall x\, \forall y\, (((x, y) \in A \times B) \sim \sim ((x \in A) \wedge (y \in B)))$. **4.** $\forall A\, \exists B\, \forall C\, ((C \subseteq A) \to (C \in B))$. **5.** Розглянемо такі предикати: $P(C)$ — "C — відповідність між A і B" та V, F, S, I і B — предикати, що відповідають властивостям всюди визначеності, функціональності, сюр'єктивності, ін'єктивності та бієктивності. (а) $\forall C\, (P(C) \sim (C \subseteq A \times B))$. (б) $\forall C\, (V(C) \sim P(C) \wedge \mathrm{Pr}_1 C = A)$. (в) $\forall C\, (F(C) \sim P(C) \wedge ((a, b) \in C \to \neg((a, d) \in C \wedge \neg(b = d))))$. (г), (д) Див. (б) і (в). (е) $\forall C\, (B(C) \sim P(C) \wedge V(C) \wedge F(C) \wedge S(C) \wedge I(C)))$. **6.** Позначимо через $V(R)$ предикат "R — відношення на M", через $P(R)$ — "R — рефлексивне відношення", через $A(R)$ — "R — антисиметричне відношення". (а) $\forall R\, (V(R) \sim (R \subseteq M \times M))$. (б) $\forall R\, (P(R) \sim (V(R) \wedge (i_M \subseteq R)))$. (в) $\forall R\, (A(R) \sim (V(R) \wedge (((a, b) \in R \wedge (b, a) \in R) \to (a = b))))$. 7. (а) "$a$ — верхня грань A" $\sim \forall x\, (x \in A \to x \le a))$. (в) "$a$ — точна верхня грань A" $\sim \sim (V(a) \wedge \forall x\, (V(x) \to (a \le x)))$, де $V(x)$ — предикат "x — верхня грань A"

(див. (а)). (д) "a — найбільший елемент A" $\sim ((a \in A) \wedge \forall x\, (x \in A \to x \le a))$ або рівносильне означення $(a \in A) \wedge V(a)$ (див. (в)). (є) "a — максимальний елемент A" $\sim ((a \in A) \wedge (\neg(\exists x\, ((x \in A) \wedge (a < x)))))$. **8.** (б), (в) — істинні. **9.** (в), (г) — істинні. Властивість щільності числової множини. **10.** (а) $\forall x\, (x \in (A \backslash B) \cap C \sim (x \in A \backslash B \wedge x \in C) \sim (x \in A \wedge \neg(x \in B) \wedge x \in C) \sim \sim ((x \in A \wedge x \in C \wedge \neg(x \in B)) \vee (x \in A \wedge x \in C \wedge \neg(x \in C))) \sim (x \in A \wedge x \in C \wedge \wedge (\neg(x \in B) \vee \neg(x \in C))) \sim (x \in A \cap C \wedge \neg(x \in B \cap C)) \sim x \in (A \cap C) \backslash (B \cap C))$. **11.** Покладемо $a = 4$, $b = 5$ і $x = 0$. Тоді з $ax = bx = 0$ виведемо $a = b$ або $4 = 5$, тобто $2 \cdot 2 = 5$.

4. Теорія графів

4.1. 1. (а) $A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$;

(б) $A = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$.

A — матриця суміжності, B — матриця інцидентності. **2.** (а) $E = \{(a, b), (a, c), (a, e), (b, d), (b, e), (c, e)\}$. (б) $E = \{(a, b), (a, e), (b, c), (c, d), (d, e)\}$. **3.** (а) $V = \{1, 2, 3, 4, 5\}$, $E = \{(1, 2), (1, 5), (2, 3)\}$. (б) $V = \{A, B, C, D, E, F\}$, $E = \{(A, D), (B, D), (C, E), (C, F)\}$. **4.** (а) Порядок матриці. (б) Половина кількості одиничних елементів матриці A. (в) Визначити множини V і E та утворити матрицю B (див. 4.1.1). **5.** Розглянемо елемент c_{ij} матриці BB^{T}: $c_{ij} = \sum_k b_{ik} b'_{kj}$, де $b_{ik} = 1$ та $b'_{kj} = b_{jk} = 1$ тоді й тільки тоді, коли ребро з номером k інцидентне вершинам i та j. Оскільки кожне ребро в графі інцидентне лише двом вершинам, то $c_{ij} = a_{ij}$ для $i \neq j$. Для всіх i c_{ii} дорівнюватиме числу ребер, інцидентних вершині з номером i, тому для отримання матриці A ці елементи слід замінити нулями. **6.** Нехай $G = (V, E)$ — граф із шістьма вершинами. Для довільної вершини $v \in V$ у графі G має місце одна з двох ситуацій: або 1) v суміжна деяким трьом вершинам w_1, w_2, w_3; або 2) v несуміжна з жодною з вершин деякої трійки вершин u_1, u_2, u_3. Якщо в першому випадку якісь дві з вершин w_1, w_2, w_3 суміжні, то разом із v вони утворюють шукану трійку попарно суміжних вершин. Якщо ж жодні дві з цих вершин несуміжні, то вони є шуканою трійкою попарно несуміжних вершин. Аналогічні міркування проводяться і для другої ситуації.

4.2. 2. $n(n-1)/2$. **3.** (а), (в), (г) — так. **4.** $n(n-1)/2 - k$ (див. 4.2.2). **5.** За означеннями операцій об'єднання та доповнення графів $G \cup \overline{G} = (V, V^{(2)})$, тобто є повним графом. **6.** Якщо розглядати елементи матриць суміжності графів як булеві константи, то матриці для графів (а) і (б) можна отримати як результат відповідно поелементної диз'юнкції і кон'юнкції матриць A_1 і A_2, а матрицю графа (г) — як заперечення всіх недіагональних елементів матриці A_1. **7.** Шуканою біекцією є відповідність $\psi = \{((v, w), (\varphi(v), \varphi(w))) \mid (v, w) \in E_1, (\varphi(v), \varphi(w)) \in E_2\}$. **8.** Однакова кількість вершин є наслідком взаємно однозначної відповідності між множинами вершин, а однакова кількість ребер — наслідком біекції між множинами ребер (див. попередню задачу). **9.** Справедливість твердження випливає з таких властивостей біекції: 1) i_V — біекція; 2) якщо f — біекція, то f^{-1} — біекція; 3) якщо f і g відповідні біекції, то $f \circ g$ — біекція. **10.** З означення доповнення маємо: $(v, w) \in \overline{E}$ тоді й тільки тоді, коли $(v, w) \notin E$. **11.** Граф із задачі 4.1.1(в). **12.** Якщо m — кількість ребер самодоповнювального графа $G = (V, E)$ з n вершинами, то $n(n-1)/2 - m = m$ (див. 4.2.4 і 4.2.8), тобто $n(n-1) = 4m$. Із того, що число $n(n-1)$ кратне 4, випливає, що $n = 4k$ або $n = 4k + 1$, $k \in N$. **13.** Із попередньої задачі маємо, що кількість ребер самодоповнювального графа G з n вершинами дорівнює $m = n(n-1)/4$ і $n = 4k$ або $n = 4k + 1$. Звідси отримаємо $m = 4k^2 - k$ або $m = 4k^2 + k$, $k \in N$. **14.** $G = (V, E)$, де $V = \{1, 2, 3, 4, 5, 6\}$, $E = \{(1, 2), (1, 3), (2, 3), (2, 5), (3, 4), (5, 6)\}$.

4.3. 1. $n - 1$. **2.** Степінь вершини v з номером i дорівнює сумі елементів i-го рядка матриці A або матриці B, бо ця сума визначає кількість вершин (для A) чи кількість ребер (для B), відповідно суміжних

або інцидентних v. **3.** Елемент $a_{ii}^{(2)}$ матриці A^2 дорівнює $a_{ii}^{(2)} = \sum_k a_{ik}a_{ki} = \sum_k a_{ik}^2 = \sum_k a_{ik}$, бо $a_{ik} = a_{ki}$ і $a_{ij} \in \{0, 1\}$. Отже, $a_{ii}^{(2)} = \delta(v_i)$ (див. попередню задачу). **4.** Див. 4.1.5. **5.** Із формули (4.1) випливає, що для кубічного графа з n вершинами і m ребрами виконується рівність $3n = 2m$. Отже, n — кратне 2 і m — кратне 3. **6.** Припустімо, що такий кубічний граф G з n вершинами існує. Тоді з того, що степені всіх вершин графа G та його доповнення дорівнюють 3, випливає, що $(n-1) - 3 = 3$, тобто $n = 7$. Звідси $\sum_i \delta(v_i) = 3 \cdot 7 = 21$, що суперечить твердженню теореми 4.1. **7.** Припустімо, що існує граф G з n вершинами, усі степені вершин якого попарно різні. Ці степені можуть дорівнювати 0, 1, 2, ..., $n - 1$. Однак якщо в графі G є вершина степеня 0 (ізольована), то в ньому не може бути вершини степеня $n - 1$. Отже, степені принаймні двох вершин збігаються. **8.** Ні. Якщо припустити, що такий граф G існує та що в ньому є тільки дві вершини v і w з однаковими степенями, то для $\delta(v) = \delta(w) = 0$ матимемо, що степені решти $n - 2$ вершин набувають усіх значень із множини $\{0, 1, 2, ..., n - 3\}$, а для випадку $\delta(v) = \delta(w) = n - 1$ — відповідно із множини $\{2, 3, ..., n - 1\}$. **9.** Футбольному турніру поставимо у відповідність граф $G = (V, E)$, у якому V — це множина команд, а $(v, w) \in E$ тоді й тільки тоді, коли команди v і w зіграли між собою. Степенем вершини буде кількість матчів, зіграних відповідною командою. Степені всіх вершин даного графа не можуть бути непарними (див. теорему 4.4). **10.** Нехай графи $G_1 = (V_1, E_1)$ і $G_2 = (V_2, E_2)$ ізоморфні, а φ — відповідне взаємно однозначне відображення V_1 на V_2. Розглянемо довільну вершину $v \in V_1$. Якщо $\delta(v) = k$, то в графі G_1 існує рівно k інцидентних вершині v ребер (v, v_1), (v, v_2), ..., $(v, v_k) \in E_1$. Тоді за означенням ізоморфізму в графі G_2 для вершини $\varphi(v)$ існуватиме також рівно k інцидентних їй ребер $(\varphi(v), \varphi(v_1))$, $(\varphi(v), \varphi(v_2))$, ..., $(\varphi(v), \varphi(v_k)) \in E_2$. Отже, $\delta(\varphi(v)) = k$.

4.4. 1. Розглянути граф $G = (V, E)$, де $V = \{u, v, w, s, t\}$, $E = \{(v, t), (t, u), (u, s), (t, s), (t, w)\}$. **2.** Використати метод доведення від супротивного та результат леми 4.2. **3.** Розглянемо деякий достатньо довгий (наприклад, максимальний) ланцюг L у даному графі, що починається з деякої вершини v. Цей ланцюг, зайшовши вперше в будь-яку проміжну вершину w по деякому ребру (z, w), за умовою задачі може продовжитись по іншому ребру (w, t), інцидентному w $(t \neq z)$. Оскільки множина вершин графа скінченна, то на якомусь кроці в ланцюзі

L деяка з пройдених раніше вершин повториться. Отже, ланцюг L містить шуканий цикл. **4.** Припустімо, що в деякому зв'язному графі G два прості ланцюги L_1 і L_2 максимальної довжини не мають жодної спільної вершини. Нехай v_1 і v_2 — початкові, а w_1 і w_2 — заключні вершини ланцюгів L_1 і L_2 відповідно. Позначимо через z_1 і z_2 найближчі вершини ланцюгів L_1 і L_2, тобто вершини, що зв'язані між собою простим ланцюгом найменшої довжини. Порівнявши довжини чотирьох простих ланцюгів, які починаються з v_1 або v_2, закінчуються у w_1 або w_2 та проходять через z_1 і z_2, дістанемо суперечність умові максимальності ланцюгів L_1 і L_2. **5.** Якщо v і w незв'язані, то вони знаходяться у різних компонентах зв'язності G_1 і G_2. Тоді сума степенів усіх вершин підграфа G_1 (або G_2) непарна, що суперечить теоремі 4.3. **6.** (а) $V = \{1, 2, 3\}$, $E = \{(1, 2), (2, 3)\}$. (б) $V = \{1, 2, 3, 4\}$, $E = \{(1, 2), (2, 3), (3, 4)\}$. (в) $V = \{1, 2, 3, 4, 5, 6\}$, $E = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 5), (3, 6)\}$. (г) K_n. **7.** Перша нерівність випливає безпосередньо з означень радіуса і діаметра графа G. Назвемо простий ланцюг *діаметральним*, якщо відстань між його початком і кінцем дорівнює діаметру графа. Для обгрунтування другої нерівності слід довести, що будь-який діаметральний ланцюг у графі G проходить через центральну вершину. **8.** Із нерівності $D(G) \geq 3$ випливає, що в графі G існує простий ланцюг $p, (p, r), r, (r, s), s, (s, t), t$ такий, що вершини p і t, p і s та r і t несуміжні. Проаналізувавши всі можливі пари вершин v та w (подібно до того, як це зроблено в доведенні теореми 4.6), дійдемо висновку, що вони зв'язані в графі $\overline{G}$, а також, що $d(v, w) \leq 2$ для всіх вершин $v \neq r$, $w \neq s$ і $d(r, s) \leq 3$. **9.** Якщо якийсь самодоповнювальний граф G незв'язний, то за теоремою 4.6 його доповнення $\overline{G}$ — зв'язний граф, що суперечитиме умові ізоморфізму графів G та $\overline{G}$. **10.** Якщо $D(G) = 1$, то G — повний граф, а жоден нетривіальний повний граф не є самодоповнювальним. Нехай $D(G) > 3$, тоді $D(\overline{G}) \leq 3$ (див. 4.4.8). У такому разі графи G і $\overline{G}$ не можуть бути ізоморфними. Для самодоповнювального графа G із задачі 4.1.1(в) $D(G) = 3$, а для $G = C_5$ $D(G) = 2$. **11.** Див. розв'язання задачі 4.3.10. **12.** $V_1 = V_2 = \{1, 2, 3, 4, 5, 6\}$, $E_1 = \{(1, 2), (2, 3), (3, 4), (4, 5), (3, 6)\}$ і $E_2 = \{(1, 2), (2, 3), (3, 4), (4, 5), (4, 6)\}$ або $E_1 = \{(1, 2), (1, 3), (1, 6), (2, 3), (2, 4), (3, 4), (4, 5), (5, 6)\}$ і $E_2 = \{(1, 2), (1, 5), (1, 6), (2, 3), (2, 4), (3, 4), (4, 5), (5, 6)\}$.

4.5. 1. (а), (б), (в) — так. **2.** (а) За теоремою 4.9 $a_{ii}^{(3)}$ дорівнює кількості шляхів довжини 3, що починаються і закінчуються у вершині i.

Ця кількість удвічі більша від числа трикутників із вершиною i, тому що існує два способи обійти кожен із цих трикутників, починаючи і завершуючи обхід у вершині i. (б) У кожному трикутнику існує 6 різних простих циклів, що відрізняються початком і напрямком обходу, і всіх їх підраховують у сумі діагональних елементів матриці A^3 (див. (а)). **3.** Справедливість твердження випливає з теореми 4.8 і того, що відстань між вершинами є довжиною найкоротшого шляху, який їх з'єднує.

4.6. 1. Розв'язки більшості задач цього розділу див. у [16, 21, 24].

4.7. 1. Ні. Якби деякий зв'язний граф з n вершинами й $n - 1$ ребром мав цикл, то, вилучивши будь-яке ребро, що належить цьому циклу, ми отримали б зв'язний граф з n вершинами й $n - 2$ ребрами (див. теорему 4.7). Останній висновок суперечить наслідку 4.8.1. **2.** Позначимо кількість кінцевих вершин у даному дереві через k. Тоді $2(n - 1) = \sum\delta(v) \geq 3(n - k) + k$, звідки випливає $k \geq n/2 + 1$. **3.** Позначимо кількість ребер графа G через m. З умови випливає, що $m \leq n$. Рівність $m = n$ призводить до суперечності, бо, з одного боку, $\sum\delta(v) = 2m$ (теорема 4.3), а з іншого — $\sum\delta(v) = m$ (усі вершини кінцеві). Якщо $m = n - 1$, то G — дерево, у якому єдину некінцеву вершину з'єднано $n - 1$ ребром з усіма іншими. Якщо ж $m < n - 1$, то граф G незв'язний (див. наслідок 4.8.1). **4.** Рівність $n(n - 1)/2 - (n - 1) = n - 1$ (див. 4.2.4) виконується для $n = 1$ або $n = 4$, що відповідає тривіальному графу або графу із задачі 4.1.1(в). **5.** Див. попередню задачу та задачу 4.2.11. **6.** Алгоритм побудови кістякового дерева для зв'язного графа описано в наслідку 4.11.5. Навпаки, якщо граф $G = (V, E)$ має кістякове дерево $T = (V, E_T)$, то зв'язність G випливає зі зв'язності T і включення $E_T \subseteq E$. **7.** Це твердження є наслідком нижньої оцінки для кількості ребер графа G (див. теорему 4.8). **8.** (а) Якщо G — ліс, то $\nu(G) = 0$ (див. наслідок 4.11.3). Нехай граф G з n вершинами має k компонент зв'язності і $\nu(G) = 0$. Тоді кількість його ребер $m = n - k$, тобто збігається з нижньою оцінкою для числа ребер (див. теорему 4.8). Такий граф ациклічний (див. розв'язання задачі 4.7.1). (б) Нехай граф G має тільки один простий цикл. Тоді після вилучення будь-якого ребра цього циклу отримаємо ациклічний граф G' із тією самою кількістю компонент зв'язності. За результатом попереднього пункту $\nu(G') = 0$. Отже, $\nu(G) = 1$. Нехай для графа G з n вершинами та k компонентами зв'язності виконується $\nu(G) = 1$. Тоді кількість його ребер $m = n - k + 1$, і за наслідком 4.11.4

у G є принаймні один цикл Z. Після вилучення ребра з циклу Z графа G отримаємо граф G' з n вершинами, k компонентами зв'язності й $n - k$ ребрами. Оскільки $\nu(G') = 0$, то G' — ациклічний граф. Отже, цикл Z — єдиний у графі G. З єдиності Z випливає його простота. (в) Це твердження можна довести методом математичної індукції за кількістю циклів у графі G. **9.** nm. **10.** Дерево T є ациклічним графом, тому не містить жодного циклу непарної довжини. Отже, за наслідком 4.12.1 теореми Кеніга T — двочастковий граф. Повними двочастковими графами є дерева виду $K_{1,n}$ або $K_{n,1}$, $n \geq 1$. Для всіх інших повних двочасткових графів $K_{n,m}$ ($n > 1$, $m > 1$) маємо $nm \geq n + m$, тобто кількість ребер у них не менша числа вершин, тому вони мають цикл (див. наслідок 4.11.4). **11.** Пряма сума двох повних графів K_n і K_m. **12.** Так, бо кожне натуральне число k можна подати у вигляді добутку двох натуральних множників, наприклад $k = 1 \cdot k$ (див. 4.7.9). **13.** Найменша — $n - 1$, найбільша — $n^2/4$ для парного n і $(n^2 - 1)/4$ для непарного n. **14.** Див. результат попередньої задачі. **15.** Якщо $n > 1$ і $m > 1$, то діаметр і радіус збігаються і дорівнюють 2. Якщо одна з часток складається з однієї вершини, а інша — більше ніж з однієї, то діаметр дорівнює 2, а радіус — 1. Нарешті, $D(K_{1,1}) = R(K_{1,1}) = 1$. **16.** Для доведення твердження слід використати результати теорем 4.9 і 4.12.

4.8. 1. Якщо з плоскої карти вилучити будь-яке ребро, то отримаємо плоску карту. **2.** Від однієї (вершина лежить усередині грані) до k граней (вершина належить межам k різних граней). **3.** Доведення можна провести методом математичної індукції за кількістю вершин дерева. Одна грань — зовнішня. **4.** $2(n - 1)$ (див. теорему 4.11 і лему 4.5). **5.** K_2. **6.** (а) Ні, бо для нього не виконується твердження наслідку 4.13.2. (б) Так. Тріангуляція. **7.** Ні. Для обґрунтування використати формулу Ейлера (4.3) і нерівність із наслідку 4.13.2. **8.** Див., наприклад, 4.8.6(б). **9.** Якщо плоский граф не містить трикутників, то $\Delta_r \geq 4$ для всіх $r \in P$. Шукану нерівність отримаємо зі співвідношень $2m = \sum_r \Delta_r \geq 4|P| = 4(m - n + 2)$. **10.** Припустімо супротивне. Тоді отримаємо несумісну систему нерівностей: $2m = \sum_v \delta(v) \geq 3n$ і $2m = \sum_r \Delta_r \geq$ $\geq 6|P| = 6(m - n + 2)$. **11.** Припустімо, що існує максимальний плоский граф G, який не є тріангуляцією. Тоді в ньому є грань Г, обмежена не трикутником, а простим циклом, який містить не менше чотирьох вершин $v_1, v_2, v_3, v_4, \ldots$. За означеннями грані і плоского графа неможливо, щоб ребра (v_1, v_3) та (v_2, v_4) одночасно належали G. Залежно від того, чи є Г внутрішньою або зовнішньою гранню,

відсутнє зі згаданої пари ребро можна провести в графі G так, що дістанемо плоский граф. Це суперечить максимальності графа G. Навпаки, нехай G — тріангуляція. Оскільки граф G зв'язний, то будь-яке ребро, після додавання якого G залишається плоским, має розбивати деяку грань графа G на дві грані. Однак у тріангуляції всі грані обмежено трикутниками, і жодну з них не можна розбити на дві шляхом додавання ребра. **12.** За лемою 4.5 для тріангуляції матимемо $3|P| = 2|E|$. Використовуючи цю рівність, із формули Ейлера (4.3) отримаємо $|E| = 3n - 6$ і $|P| = 2n - 4$.

4.9. 1. (а) n. (б)–(г), (е) 2. (д) 3. **2.** $n - 1$. **3.** Можна довести, що граф біхроматичний тоді й тільки тоді, коли він двочастковий. Відтак використати теорему Кеніга. **4.** (а) Наприклад, незв'язний граф G_1 складається з двох трикутників C_3, а $G_2 = C_6$. (б) $G_1 = K_4$, а G_2 — пряма сума всіх простих циклів графа K_4. (в) $G_1 = (V_1, E_1)$, $V_1 = \{1, 2, 3, 4, 5, 6\}$, $E_1 = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4), (3, 6), (4, 5), (5, 6)\}$ і $G_2 = (V_2, E_2)$, $V_2 = V_1$, $E_2 = \{(1, 2), (1, 6), (2, 3), (2, 5), (2, 6), (3, 4), (3, 5), (4, 5), (5, 6)\}$. **5.** Доведення проведемо, як у теоремі 4.18, індукцією за кількістю n вершин планарного графа. Для $n \leq 5$ твердження виконується. Припустимо, що хроматичне число будь-якого планарного графа з t вершинами не перевищує 5. Розглянемо довільний планарний граф G з $t + 1$ вершиною і вилучимо з нього вершину v, степінь якої не перевищує 5. Позначимо отриманий граф через G'. Нехай $S(v)$ — множина синів вершини v. Якщо $|S(v)| \leq 4$, то, розфарбувавши за припущенням індукції вершини G' у п'ять кольорів і повернувши вершину v на місце, зможемо пофарбувати її в один із п'яти кольорів, який не було використано при розфарбуванні вершин із $S(v)$. Якщо ж $|S(v)| = 5$, то в множині $S(v)$ є принаймні дві несуміжні вершини v_1 і v_2 (в іншому разі граф G містив би K_5 і не був планарним). Виконаємо в графі G' ототожнення (або злиття) вершин v_1 і v_2 у вершину w. Отримаємо плоский граф G'', який за індуктивним припущенням розфарбуємо у п'ять кольорів, і зафіксуємо це розфарбування. У графі G вершини v_1 і v_2 пофарбуємо в колір вершини w, а вершину v — у колір, який не було використано у розфарбуванні вершин з $S(v)$. Так дістанемо шукане правильне розфарбування графа G п'ятьма фарбами. **6.** $\chi(K_n) = n$, а після вилучення довільної вершини з K_n отримуємо повний граф K_{n-1}, хроматичне число якого дорівнює $n - 1$. **7.** Припустимо, що якийсь критичний граф

G незв'язний і G_1, G_2, …, G_k — його компоненти зв'язності. Вилучимо довільну вершину з тієї компоненти G_i, хроматичне число якої не перевищує хроматичні числа решти компонент зв'язності. Отримаємо суперечність, бо після такого вилучення хроматичне число графа G не зміниться. **8.** Єдиний 2-критичний граф — K_2. Усі 3-критичні — це C_n, де $n = 2k + 1$, $k \in N$.

4.10. 1. Необхідність твердження випливає з того, що будь-який цикл у графі можна зобразити у вигляді об'єднання простих циклів, які містяться в ньому (див. лему 4.2). Припустимо, що граф G можна подати у вигляді об'єднання простих циклів і Z_1 — один із них. Якщо Z_1 містить усі ребра графа G, то G — ейлерів граф. Відтак повторимо міркування, наведені в доведенні достатності теореми 4.20. **2.** Кожна вершина є вершиною деякого ребра, а кожне ребро даного графа належить його ейлеровому циклу. **3.** n — непарне. **4.** Для парних n і m. **5.** Доведення даного твердження подібне до доведення теореми 4.20. **6.** "Подвоїмо" кожне ребро даного графа G, перетворивши його в мультиграф G' (див. розділ 4.13). Степені всіх вершин G' парні, тому за теоремою Ейлера в G' існує цикл, що містить усі його ребра. Цьому циклу відповідає шуканий циклічний маршрут у графі G. **7.** Граф, для якого існує такий маршрут, має бути ейлеровим (див. розв'язання попередньої задачі). **8.** Нехай множиною вершин повного графа K_n є $V = \{1, 2, …, n\}$. Тоді 1, (1, 2), 2, (2, 3), …, $n - 1$, $(n - 1, n)$, n, $(n, 1)$, 1 — гамільтонів цикл. **9.** Припустимо, що даний граф G має гамільтонів цикл Z. Тоді окрім двох вершин v і w степеня 3 всі інші вершини u_1, u_2, …, u_n повинні мати степінь 2, тому що граф, у якому є кінцева вершина, не є гамільтоновим. Гамільтонів цикл Z графа G міститиме усі ребра, інцидентні вершинам u_1, u_2, …, u_k, і по два з трьох ребер, інцидентних v та w. Кожне з ребер, що не потрапило в Z, не може з'єднувати вершини v і w (за умовою вони несуміжні). Крім того, ці ребра не можуть бути інцидентні жодній із вершин u_1, u_2, …, u_k (їх степені дорівнюють 2). Прийшли до суперечності. Отже, G не може мати гамільтонового циклу. **10.** Граф, що складається з двох (або більше) трикутників з єдиною спільною вершиною, є ейлеровим, але не є гамільтоновим. K_4 — гамільтонів, але не ейлерів.

4.11. 2. У сумі (а) підраховано кількість початків усіх дуг орграфа G. Ця кількість дорівнює числу дуг $|E|$, а також числу кінців усіх дуг орграфа G, яке обчислюється у (б). **3.** Так, якщо в орграфі дозволено

петлі, в іншому разі — ні. **4.** Твердження можна довести методом математичної індукції за числом вершин повного орграфа. **5.** Повний орграф G може не мати жодної недосяжної (тупикової) вершини, якщо він гамільтонів. Він може мати тільки одну недосяжну (тупикову) вершину, і вона буде єдиним джерелом (стоком) орграфа G (див. попередню задачу). Припущення про існування більше ніж двох недосяжних (тупикових) вершин у повному орграфі призведе до суперечності з результатом попередньої задачі. **6.** (а) Кількість дуг у повному орграфі G збігається з числом ребер у відповідному неорієнтованому графі G', тобто у графі, який можна отримати з G, замінивши в ньому всі орієнтовані дуги на неорієнтовані ребра. Другу суму, розкривши дужки, можна перетворити до вигляду $n(n-1) - n(n-1)/2$. З іншого боку, цю суму можна розглядати як суму нівстепенів заходу всіх вершин орграфа G, бо в повному орграфі $n - 1 - \delta^+(v_i) = \delta^-(v_i)$, $i = 1, 2, 3, …, n$. (б) Рівності можна довести методом математичної індукції за числом вершин повного орграфа. **7.** Твердження доведемо методом математичної індукції за числом n вершин повного орграфа G. Неважко переконатись у справедливості цього твердження для $n = 2$ і $n = 3$. Припустимо, що це твердження виконується для всіх повних орграфів із k вершинами і розглянемо довільний турнір $G = (V, E)$ з $k + 1$ вершиною. Вилучимо будь-яку вершину v з орграфа G. Отримаємо повний орграф G', у якому за припущенням індукції існує гамільтонів ланцюг L. Нехай u_1, u_2, …, u_k — послідовність вершин, через які проходить ланцюг L. Якщо $(v, u_1) \in E$, то v, u_1, u_2, …, u_k — послідовність вершин шуканого ланцюга для турніру G. Якщо ж $(u_1, v) \in E$, то розглянемо першу вершину u_i ланцюга L, для якої $(v, u_i) \in E$ (якщо така вершина існує). Тоді $(u_{i-1}, v) \in E$, і u_1, u_2, …, u_{i-1}, v, u_i, …, u_k — послідовність вершин гамільтонового ланцюга. Якщо такої вершини u_i не існує, то шуканою послідовністю вершин буде u_1, u_2, …, u_k, v. **8.** Необхідність. Нехай G — сильно зв'язний орграф і Z — його замкнений маршрут, що проходить через максимально можливе число вершин u_1, u_2, …, u_k, u_1. Якщо цей маршрут не кістяковий, то розглянемо вершину v, яка не потрапила до нього. Оскільки G сильно зв'язний, то існують маршрути Z_1 і Z_2, які складаються з послідовностей вершин u_1, w_1, w_2, …, w_s, v та v, y_1, y_2, …, y_t, u_1 відповідно. Тоді послідовність u_1, u_2, …, u_k, u_1, w_1, w_2, …, w_s, v, y_1, y_2, …, y_t, u_1 відповідає замкненому маршруту, який містить більше число вершин, ніж у Z, що суперечить припущенню про максимальність Z. Отже, Z — кістяковий маршрут. Достатність випливає з того, що наявність в орграфі G замкненого кістякового маршруту гарантує існування шляху, який веде з будь-якої вершини v в будь-яку іншу вершину w орграфа G. **9, 10.** Див. розв'язання задачі 4.11.8. **11.** Неважко довести, що ациклічний орграф G завжди має принаймні один стік. Нехай u_1 — стік даного орграфа G. Покладемо $f(u_1) = n$ і вилучимо з G вершину u_1. Отримаємо ациклічний орграф G_1, в якому існуватиме стік u_2. Нехай $f(u_2) = n - 1$ і вилучимо u_2 з G_1. Продовжуючи цю процедуру, отримаємо шукану правильну нумерацію. **12.** Слід довести, що транзитивний повний орграф є ациклічним і скористатись результатом попередньої задачі. **13.** Нижня оцінка випливає з того, що сильно зв'язний орграф G має замкнений кістяковий маршрут (див. 4.11.8), кількість дуг у якому не менша від кількості вершин орграфа G. Із кожної з n вершин орграфа G можна провести $n - 1$ дугу в усі інші його вершини, тому максимальне число дуг у G дорівнює $n(n - 1)$. **14.** Для доведення твердження використаємо результати задач 4.11.7 і 4.11.8. У будь-якому повному орграфі $G = (V, E)$ існує гамільтонів ланцюг L. Нехай вершина u_1 — початок, а вершина u_k — кінець ланцюга L. Якщо $(u_k, u_1) \in E$, то, додавши (u_k, u_1), u_1 до L, отримаємо замкнений кістяковий маршрут в орграфі G. В іншому разі змінимо спочатку орієнтацію дуги (u_1, u_k), і тоді додамо пару (u_k, u_1), u_1 до L. **15.** Дане твердження можна довести так само, як теорему 4.9. **16.** Доведення даного твердження подібне до доведення теореми 4.20. **17.** Див. задачі 4.11.7 і 4.11.8.

5. Комбінаторика

5.1. 1. (б) $|\varnothing|$, $|A \setminus B|$, $|A \,\Delta\, B|$, $|A \cup B|$, $|A| + |B|$. **2.** Розглянути випадки $A \cap B = \varnothing$ і $A \cap B \neq \varnothing$. **3.** (а), (б) Однакова кількість, бо між сукупностями даних підмножин можна встановити бієкцію. Наприклад, для (а) такою бієкцією буде відповідність $\{(C, A \setminus C) \mid C \subseteq A\}$. (в) Тих, що не включають множину B. **4.** (а) 6. (б) 10. **5.** (а) 80 %. (б) 20 %. (в) 60 %. (г) 70 %. **6.** (а) Позначимо через A_k множину чотиризначних натуральних чисел, що діляться на k. Загальна кількість чотиризначних натуральних чисел дорівнює 9000, тому шукану кількість можна обчислити за формулою $9000 - |A_3 \cup A_5 \cup A_7|$. Крім того, $|A_k| = [9999/k] - [999/k]$ і $A_k \cap A_l = A_m$, де m — найменше спільне кратне чисел k і l. Для обчислення суми даних чисел слід від суми всіх чотиризначних натуральних чисел відняти суму чисел, що діляться на 3, 5, або 7.

Кількість — 4114, сума — 22624427. (б) 6600 і 36306000 (див. (а)). **7.** 46. Обчислити значення $199 - |A_2 \cup A_3 \cup A_5 \cup A_7 \cup A_{11} \cup A_{13}|$ (див. попередню задачу). **8.** Позначимо відповідні множини студентів через Д, М, А і П. Тоді $|Д \cap М| = |Д| + |М| - |Д \cup М| \geq 70\,\% + 75\,\% - 100\,\% = 45\,\%$ і $|А \cap П| \geq 65\,\%$. Отже, для кількості студентів, що склали всі чотири іспити, матимемо $|Д \cap М \cap А \cap П| = |Д \cap М| + |А \cap П| - |(Д \cap М) \cup \cup (А \cap П)| \geq 45\,\% + 65\,\% - 100\,\% = 10\,\%$. **9.** $33^3 \cdot 10^4$. **10.** 8!. Є вісім способів розмістити першу туру на першій вертикалі шахівниці, сім способів розмістити другу туру на другій вертикалі і т. д., один спосіб розмістити восьму туру на останній вертикалі. **11.** $mn(m - 1)(n - 1)$. Першу туру можна поставити на одну з mn клітинок шахівниці, а другу — на будь-яку з решти $(m - 1)(n - 1)$ клітинок, що не знаходяться на одній вертикалі і на одній горизонталі з клітинкою, яку займає перша тура. **12.** (а) $9 \cdot 10^{n-2}$. Перші дві цифри можна обрати дев'ятьма способами, а для вибору кожної з решти $n - 2$ цифр є 10 варіантів. (б) 9^n. (в) 5^n. (г) $9 \cdot 10^{n-1} - 5^n$. Від числа всіх n-значних десяткових чисел слід відняти кількість n-значних чисел, усі цифри яких непарні. (д) $8 \cdot 9^{n-1}$. (е) $9 \cdot 10^{n-1} - 8 \cdot 9^{n-1}$ (див. (г) і (д)). **13.** 3^n. Розглянемо таблицю, два стовпчики якої відповідають підмножинам X і Y множини A, а рядки — елементам множини A. У клітинку таблиці записуємо 1, якщо відповідний елемент належить відповідній із підмножин (X або Y), 0 — в іншому разі. Ситуації, коли підмножини X та Y не перетинаються, відповідає таблиця, в якій у кожному рядку записано одну з трьох пар: (0, 0), (0, 1) або (1, 0). **14.** $(a_1 + 1)(a_2 + 1) \cdots (a_r + 1)$. Будь-який дільник числа n можна подати у вигляді $p_1^{b_1} p_2^{b_2} \ldots p_r^{b_r}$, де $b_i = 0, 1, 2, \ldots, a_i$, $i = 1, 2, \ldots, r$. **15.** Кожним чотирьом вершинам n-кутника взаємно однозначно відповідає пара діагоналей, що перетинаються в одній точці всередині n-кутника. Отже, шукана кількість збігається з кількістю невпорядкованих четвірок вершин, яка дорівнює $n(n - 1)(n - 2)(n - 3)/4!$. **16.** $mn(n - 1)/2 + mn(m - 1)/2 = mn(m + n - 2)/2$. **17.** (а) $4 \cdot 5^{n-1}$. (б) Якщо n — парне, то $3 \cdot 4^{n/2}$; якщо n — непарне, то $3^2 \cdot 4^{(n-1)/2}$. (в) $4^2 \cdot 5^{n-2}$. **18.** (а) Оскільки відповідність C між A і B — це довільна підмножина декартового добутку $A \times B$, то кількість таких відповідностей дорівнює кількості елементів булеана множини $A \times B$, яка містить nm елементів. За теоремою 1.1 це число дорівнює 2^{nm}. (б) Будь-яку відповідність C між A і B можна однозначно задати множиною образів $C(a)$, $a \in A$. Для всюди визначеної відповідності C кожна з цих множин (що є підмножиною множини B) має бути непорожньою. Кількість непорожніх підмножин множини B дорівнює $2^m - 1$. За правилом множення кількість всюди визначених відповідностей дорівнюватиме $(2^m - 1)^n$. (в) Множиною образів $f(a)$ функціональної відповідності f між A і B для кожного $a \in A$ може бути або деяка одноелементна підмножина множини B, або порожня множина. Отже, є $m + 1$ варіант обрати образ для кожного з n елементів множини A, тому кількість функціональних відповідностей дорівнює $(m + 1)^n$. (г) $(2^n - 1)^m$ (див. (б)). (д) $(n + 1)^m$ (див. (в)). (е) $n!$, якщо $n = m$, 0 — в іншому разі. **19.** Нехай $m = n^2$, $k = n^2 - n$, $l = k/2$. (а) 2^m. Кількість відношень на множині M дорівнює числу елементів булеана множини $M \times M$, яка містить $m = n^2$ елементів. (б) 2^k. Будь-яке рефлексивне відношення R на множині M можна подати у вигляді $R = i_M \cup R'$, де R' — множина недіагональних елементів із $M \times M$. Кількість варіантів обрати R' дорівнює кількості елементів булеана множини $(M \times M) \setminus i_M$, що містить $k = n^2 - n$ елементів. (в) Це число дорівнює різниці кількостей усіх відношень і рефлексивних відношень на M, тобто дорівнює $2^m - 2^k$. (г) 2^k (див. (б)). (д) $2^m - 2^k$ (див. (в) і (г)). (е) Оскільки жодне відношення на M не може бути одночасно рефлексивним і антирефлексивним, то шукане число дорівнює різниці кількості всіх відношень і суми кількостей рефлексивних і антирефлексивних відношень: $2^m - 2 \cdot 2^k = (2^n - 2)2^k$ (див. (б) і (г)). (є) Будь-яке симетричне відношення на M можна однозначно задати, вибравши довільну підмножину діагоналі i_M і довільну підмножину з елементів, розміщених над (або під) діагоналлю. Кількість елементів, з яких здійснюється вибір, дорівнює $n + l = n(n + 1)/2$, тому число симетричних відношень — 2^{n+l}. (ж) $2^m - 2^{n+l}$ (див. (в) і (є)). (з) Для того щоб утворити антисиметричне відношення R, слід узяти довільну підмножину діагоналі i_M (2^n способів) і додати до неї деяку множину недіагональних елементів, ураховуючи те, що з пари (a, b) та (b, a) симетричних недіагональних елементів до R може входити або лише один із цих елементів, або не входити жоден (3^l способів, бо кількість таких пар дорівнює l). Отже, існує $2^n \cdot 3^l$ антисиметричних відношень на множині M. (и) 2^l. (і) $2^k - 2^l = 2^l(2^l - 1)$. (ї) 3^l (див. (з)). (й) 2^l. (к) $2^l(2^l - 1)$. (л) 3^l (див. (з)).

5.2. 1. $C(n, k)$. **2.** $C(6, 2)$Є$(14, 5)$. **3.** (а) $C(50, 3) + 50C(50, 2)$. Сума трьох чисел парна, якщо або всі вони парні, або два з них непарні й одне парне. (б) $C(50, 3) + 50C(50, 2)$. (в) $2C(33, 3) + C(34, 3) + 33^2 \cdot 34$. **4.** $C(n, p) + C(n, p + 1) + \ldots + C(n, n)$. **5.** $8!$. **6.** $n!m!C(n, m)$. Розсадимо спочатку за столом усіх n чоловіків. Існує n місць між ними, які можуть займати жінки, щоб не сидіти поруч. m жінок можуть зайняти ці

місця $C(n, m)$ способами. Для кожного розташування існує $n!$ способів інакше розсадити чоловіків і $m!$ способів — жінок. **7.** $C(n + 1, k)$. Запишемо спочатку всі n нулів. Існує $n + 1$ місце (одне перед першим нулем, $n - 1$ — між нулями й одне — після останнього нуля), на які можна записувати по одній одиниці, щоб вони не стояли поруч. Отже, для кожного заданого розміщення слід обрати k місць із $n + 1$ можливих. **8.** Нехай $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_m\}$, $a_i < a_{i+1}$, $i = 1, 2, \ldots, n - 1$, $b_j < b_{j+1}$, $j = 1, 2, \ldots, m - 1$, $z = (b_{j_1}, b_{j_2}, \ldots, b_{j_n})$ — вектор значень певного монотонного відображення f, тобто $b_{j_k} = f(a_k)$, $b_{j_k} \in B$ та $j_k \le j_{k+1}$, $k = 1, 2, \ldots, n - 1$. Вектору z поставимо у відповідність двійковий кортеж за таким правилом: спочатку запишемо $j_1 - 1$ одиницю, за ними — 0, після цього нуля запишемо $j_2 - j_1$ одиниць, знову 0 і т. д., після останнього n-го нуля запишемо $m - j_n$ одиниць. Неважко переконатися, що описана відповідність взаємно однозначна, а кожен із кортежів містить n нулів і $m - 1$ одиницю. Отже, шукане число збігається з числом таких кортежів і дорівнює $C(n + m - 1, m - 1)$. **9.** $28!/(7!)^4$. **10.** (а) $(n + m + k)!/(n!m!k!)$. (б) Розглянемо довільне слово w, утворене з усіх символів a і b. Кількість таких слів — $(n + m)!/(n!m!)$, а довжина кожного з них — $m + n$. У слові w $n + m + 1$ позиція, на які можна записувати по одному символу c, щоб вони не опинилися поруч. Це можна зробити $C(n + m + 1, k)$ способами. Отже, шукана кількість — $C(n + m + 1, k)(n + m)!/(n!m!)$.

5.3. 1. (а) $16x^4 - 128x^3 + 384x^2 - 512x + 256$. (в) $1 - 5x^3 + 10x^6 - 10x^9 + 5x^{12} - x^{15}$. **2.** (а) 721 і –676. (б) $C(3, 3) + C(4, 3) + \ldots + C(15, 3)$ і $C(5, 5) + C(6, 5) + \ldots + C(15, 5)$. **3.** Для $m < k$ $C(k, m) + C(k + 1, m) + \ldots + C(n, m)$; для $m \ge k$ $C(m, m) + C(m + 1, m) + \ldots + C(n, m)$. **4.** (а) 0,976. (б) 1039,36. **5.** (а) $C(n, 5) = C(n, 12) \Rightarrow n = 17$. (б) $C(n, 5) - C(n, 4) = C(n, 6) - C(n, 5) \Rightarrow n = 7$ чи $n = 14$. (в) $n = 11$ або $n = 12$. Шукані натуральні числа n є розв'язками системи з двох нерівностей: $C(n, 7)2^{n-7} \cdot 3^7 > C(n, 6)2^{n-6} \cdot 3^6$ і $C(n, 7)2^{n-7} \cdot 3^7 > > C(n, 8)2^{n-8} \cdot 3^8$. **6.** (а) 9. (б) 10. (в) 4. (г) 13. **7.** (а) $n = 18$, $m = 8$. (б) $n = 6$, $m = 3$. **8.** (г) Ліва частина є розкладом бінома $(1 + 2)^n$. (д) Використавши рівність $(n + 1)/(k + 1)C(n, k) = C(n + 1, k + 1)$, або $C(n, k)/(k + 1) = C(n + 1, k + 1)/(n + 1)$, перетворимо ліву частину до вигляду $C(n + 1, 1)/(n + 1) + C(n + 1, 2)/(n + 1) + \ldots + C(n + 1, n + 1)/(n + 1)$ і застосуємо тотожність 3 з (5.8). (е) Використавши рівність $kC(n, k) = nC(n - 1, k - 1)$, перетворимо ліву частину до вигляду $nC(n - 1, 0) + nC(n - 1, 1) + \ldots + nC(n - 1, n - 1)$ і застосуємо тотожність 3 з (5.8).

(є) Див. (е) і тотожність 4 з (5.8). (ж) Використати тотожність 2 з (5.8). (з) Див. (е) і тотожність 6 із (5.8). Інший метод — додати тотожності (е) і (є). (и) Відняти тотожності (е) і (є). (і) Можна використати спочатку рівність із розв'язання (е), а тоді — тотожності (е) та 3 з (5.8). Інший метод використовує тотожність $(1 + x)^n = \sum_k C(n, k)x^k$. Продиференціюємо обидві її частини по x, помножимо на x і знову продиференціюємо. Відтак нідставимо в одержаній рівності замість x значення 1. **9.** Позначимо через d НСД даних чисел. Послідовно застосовуючи тотожність 2 з (5.8), отримаємо, що d — дільник числа $C(n - k, 0)$. Отже, $d = 1$. **10.** Розглянемо довільне число $C(n, k)$, розміщене всередині n-го рядка трикутника Паскаля, $k = 1, 2, \ldots, n - 1$. Із формули (5.5) маємо $k!(n - k)!C(n, k) = n!$. Оскільки права частина ділиться на n, то й ліва ділиться на n. Однак ні $k!$, ані $(n - k)!$ не можуть ділитися на n, тому з простоти n випливає, що $C(n, k)$ ділиться на n. **11.** $x^3 + y^3 + z^3 + 3x^2y + 3x^2z + 3xy^2 + 3y^2z + 3xz^2 + 3yz^2 + 6xyz$. **12.** (а) $C_9(5, 4, 0) + C_9(6, 1, 2) = 378$. (б) $C_7(2, 4, 1) + C_7(3, 1, 3) = 245$. (в) $-12C_{10}(7, 2, 1) + 16C_{10}(6, 4, 0) + 9C_{10}(8, 0, 2) = -555$. **13.** Шуканий коефіцієнт дорівнює $K = nC_2(1, 0, \ldots, 0, 1) + (n - 1)C_2(0, 1, 0, \ldots, 0, 1, 0) + 2(n - 2)C_2(0, 0, 1, 0, \ldots, 0, 1, 0, 0) + \ldots$. Останній член цієї суми залежить від n і дорівнює $k(n - k)C_2(0, 0, \ldots, 0, 1, 1, 0, \ldots, 0)$ для непарного n ($k = (n - 1)/2$) та $n^2C_2(0, 0, \ldots, 0, 2, 0, \ldots, 0)/4$ — для парного n. Обидва ці випадки можна об'єднати, якщо записати коефіцієнт K у вигляді $n + 1\cdot(n - 1) + 2(n - 2) + \ldots + (n - 2)(n - (n - 2)) + (n - 1)(n - (n - 1)) + n = 2n + n(1 + 2 + \ldots + (n - 1)) - (1^2 + 2^2 + \ldots + (n - 1)^2) = 2n + n^2(n - 1)/2 - n(n - 1)(2n - 1)/6 = n(n^2 + 11)/6$. **14.** Сума всіх коефіцієнтів розкладу многочлена дорівнює його значенню при $x_1 = x_2 = \ldots = x_k = 1$.

5.4. 1. Для розв'язання наступних задач слід використати правило множення і той факт, що розміри прямокутної сітки квадратів, у якій існує $C(n, k)$ траєкторій, що ведуть із лівої нижньої у праву верхню вершину, дорівнюють $k \times (n - k)$ або $(n - k) \times k$. (а) Усі траєкторії, що ведуть із початкової вершини $O(0, 0)$ у вершину $A(k, n - k)$ (їх кількість — $C(n, k)$), слід розбити на три групи за допомогою точок $A1(k, n - k - 2)$, $A2(k - 1, n - k - 1)$ і $A3(k - 2, n - k)$. Врахувати, що з $A2$ в A веде дві траєкторії (у той час, як з $A1$ в A та з $A3$ в A — по одній). (б) Розглянути множину T траєкторій, що ведуть з $O(0, 0)$ у вершину $A(n, n)$ (кількість таких траєкторій — $C(2n, n)$). Розбити множину T на підмножини T_k траєкторій, що ведуть з O в A і проходять через точки $A_k(k, n - k)$, $k = 0, 1, \ldots, n$. Число траєкторій, що ведуть з O в A_k, дорівнює $C(k + (n - k), k) = C(n, k)$, а з A_k в A — $C((n - k) + k, n - k) = C(n, n - k)$. Отже, за правилом множення число траєкторій у підмножині T_k дорівнює $(C(n, k))^2$. (в) Цільова вершина — $A(k, n - k)$, а вершини відповідного розбиття — $A_m(m, k - 2)$, $m = 0, 1, \ldots, n - k$. (г) Цільова вершина — $A(k, n + m - k)$, вершини розбиття — $A_m(m, n - m)$, $m = 0, 1, \ldots, k$. (д) Цільова вершина — $A(n - k, k + 1)$, вершини розбиття — $A_m(m, k)$, $m = 0, 1, \ldots, n - k$. Відповідна підмножина T_m розбиття — це множина траєкторій, які проходять через ребро, що з'єднує вершини $A_m(m, k)$ і $B_m(m, k + 1)$. Число таких траєкторій дорівнює числу траєкторій, які ведуть із початкової вершини в A_m, бо існує єдина траєкторія, яка веде з A_m в A та проходить через ребро A_mB_m. (е) Цільова вершина — $A(m - 1, n + 1)$, вершини розбиття — $A_k(k, n)$, $k = 0, 1, \ldots, m - 1$. Див. попередню задачу. (є) Цільова вершина — $A(n, m)$, вершини розбиття — $A_k(n - 1, k)$, $k = 0, 1, \ldots, m$. Див. (д). (ж) Цільова вершина — $A(n - m + k + 1, m)$. Подібно до трьох попередніх задач множину T траєкторій, які ведуть з O в A, можна розбити на підмножини T_l траєкторій, що проходять через ребро A_lB_l, де $A_l(k, l)$ і $B_l(k + 1, l)$. Число елементів множини T_l дорівнює добутку кількостей траєкторій, які ведуть з O в A_l, і тих, що ведуть із B_l в A.

5.5. 1. $C(11, 4)$. **2.** (а) Множину варіантів розподілу слід розбити на ситуації, коли порожніми є дві, одна або жодна з урн: $C(5, 2)C(14, 2) + C(5, 1)C(14, 3) + C(14, 4) = 3731$. (б) $C(14, 4)$. (в) Шукана кількість K дорівнює кількості цілочислових розв'язків рівняння $x_1 + x_2 + \ldots + x_6 = 25$, які задовольняють умови $0 \le x_i \le 5$, $i = 1, 2, \ldots, 6$. Після заміни $y_i = 6 - x_i$, $i = 1, 2, \ldots, 6$, отримаємо рівняння $y_1 + y_2 + \ldots + y_6 = 11$, кількість розв'язків якого в натуральних числах збігається з K. Отже, $K = C(10, 5)$. **3.** $C(n, m)C(k - 1, n - m - 1)$. **4.** Це число збігається з кількістю розподілів n предметів по m урнах за умови, що жодна з урн не залишається порожньою. **5.** Між множинами розв'язків даних рівнянь можна встановити таку взаємно однозначну відповідність: розв'язку $(a_1, a_2, \ldots, a_m)$ першого рівняння відповідає розв'язок $(a_1 + 1, a_2 + 1, \ldots, a_m + 1)$ другого рівняння. **6.** $C(n + m - 1, m - 1) + + C(n + m - 2, m - 1) + \ldots + C(m - 1, m - 1)$. **7.** Кількість способів збігається з числом невід'ємних цілих розв'язків рівняння $x + y + z = n$ і дорівнює $C(n + 2, 2)$. **8.** $C(n + 1, 1)C(m + 2, 2) = (n + 1)(m + 2)(m + 1)/2$. **9.** Шукана кількість способів K збігається з кількістю зображень числа n у вигляді суми k натуральних доданків, $k = 1, 2, \ldots, n$. Зображення, що відрізняються порядком доданків, вважають різними. $K = C(n-1, 0) + C(n-1, 1) + C(n-1, 2) + \ldots + C(n-1, n-1) = 2^{n-1}$. **10.** Число способів розподілити 4 кулі одного кольору по 6 пакетах дорівнює $C(9, 5)$, тому за правилом множення шукана кількість — $(C(9,5))^3$. **11.** $C(n+p-1, p-1)C(m+p-1, p-1)C(k+p-1, p-1)$, $C(n-l_1+p-2, p-2)C(m-l_2+p-2, p-2)C(k-l_3+p-2, p-2)$. **12.** $C(2n, n)$. **13.** Нехай у першій урні опинились x білих, y червоних і z синіх куль. Тоді дістанемо рівняння $x + y + z = 3n$, де $0 \le x, y, z \le 2n$. Загальна кількість розв'язків цього рівняння у невід'ємних цілих числах дорівнює $C(3n + 2, 2)$. Від неї слід відняти число розв'язків, у яких значення принаймні однієї з невідомих перевищує $2n$. Кількість розв'язків, у яких $x > 2n$, збігається із загальним числом розв'язків у невід'ємних цілих числах рівнянь $y + z = k$ для $k = 0, 1, 2, \ldots, n-1$. Це число дорівнює $C(1, 1) + C(2, 1) + \ldots + C(n, 1) = 1 + 2 + \ldots + n = n(n+1)/2$. Аналогічно можна підрахувати кількість розв'язків, коли $y > 2n$ і $z > 2n$. Отже, шукана кількість способів розподілу дорівнює $C(3n + 2, 2) - 3n(n+1)/2 = 3n^2 + 3n + 1$. **14.** Поставимо у відповідність кожному з n сортів тістечок умовну урну. Тоді будь-якому вибору k тістечок відповідатиме певний розподіл k куль по n урнах. Отже, кількість способів вибору дорівнює $C(k + n - 1, n - 1)$. **15.** Нехай a — довільний представник із заданих n груп елементів. Кількість сполук з n по m із повтореннями, які містять єдиний елемент a, збігається з кількістю сполук з $n-1$ по $m-1$ із повтореннями, тобто дорівнює $C(n+m-3, n-2)$. Аналогічно, кількість сполук з n по m із повтореннями, які містять по два елементи a, збігається з кількістю сполук з $n-1$ по $m-2$ з повтореннями і дорівнює $C(n+m-4, n-2)$ і т. д. Отже, шукана кількість — $1 \cdot C(n+m-3, n-2) + 2 \cdot C(n+m-4, n-2) + \ldots + (m-1)C(n-1, n-2) + mC(n-2, n-2) = C(n+m-1, n)$ (див. 5.3.8(ж)). Той самий результат можна отримати шляхом інших міркувань. Число усіх сполук з n по m із повтореннями дорівнює $C(n+m-1, n-1)$. Загальна кількість елементів, які входять до них, — $mC(n+m-1, n-1)$. Оскільки кожен з елементів повторюється у цій сукупності однакову кількість разів, то шукане число дорівнює $mC(n+m-1, n-1)/n = C(n+m-1, n)$.

5.6. 1. (а) 382 і 813. (б) 100 і 9930. (в) 8 і 16. **2.** (а) Покладемо $b_n = 1/a_n$. Тоді матимемо $b_n = b_{n-1} + 1$ і $b_1 = 1$. Отже, $b_n = n$ і $a_n = 1/n$. (б) Покладемо $b_n = a_n^2$. Дістанемо таке рекурентне співвідношення $b_n = 2b_{n-1} + 1$. Неважко переконатися, що для послідовності $\{b_n\}$ виконується $b_n + 1 = 2(b_{n-1} + 1)$, тобто послідовність $\{b_n + 1\}$ є геометричною прогресією зі знаменником 2 і $b_1 + 1 = 2$. Отже, $b_n + 1 = 2^n$ і $a_n = \sqrt{2^n - 1}$. **3.** Почленно віднімемо рівності $a_n = Aa_{n-1} + Ba_{n-2} + C$ і $a_{n-1} = Aa_{n-2} + Ba_{n-3} + C$. Отримаємо $a_n - a_{n-1} = A(a_{n-1} - a_{n-2}) + B(a_{n-2} - a_{n-3})$, тобто $b_n = Ab_{n-1} + Bb_{n-2}$. **4.** (а) $C_1(-3)^n + C_2(-4)^n$. (б) $C_1(2-3i)^n + C_2(2+3i)^n$. (в) $C_1(-3i)^n + C_2(3i)^n$. (г) $C_1 2^n + C_2((7-\sqrt{41})/2)^n + C_3((7+\sqrt{41})/2)^n$. (д) $(C_1 + C_2 n)(-1)^n$. (е) $C_1(1+i)^n + C_2(-1+i)^n + C_3(-1-i)^n + C_4(1-i)^n$. **5.** (а) $5 \cdot 2^n - 3^{n+1}$. (б) $3^n + i^n + (-i)^n$. (в) $3 \cdot 2^n + 2(-3)^n + 4^n$. (г) $C_1 + C_2 n + C_3(-2)^n$, де $C_1 = (14 - b - 4c)/9$, $C_2 = (b + c - 2a)/3$, $C_3 = (2b - a - c)/18$. (д) $\cos n\alpha$. (е) $a_n = \sqrt{51 \cdot 4^{n-1} - 35}$. Зробити заміну $b_n = a_n^2$, $n = 1, 2, \ldots$. **6.** (а) $n(n-1)/2 + 1$. Частинний розв'язок шукаємо у вигляді $G(n) = b_1 n^2 + b_2 n$. (б) $50 \cdot 2^n + 6{,}5 \cdot 3^n + 0{,}5 - (4n^3 + 13n^2 + 50n)$. **7.** (а) $n(n+1)(2n+1)/6$. (б) $n(n+1)(2n+1)(3n^2 + 3n - 1)/30$. **8.** Використаємо метод математичної індукції за n. База індукції: для $n = 2$ маємо співвідношення $F_{m+2} = F_1 F_m + F_2 F_{m+1} = F_m + F_{m+1}$, яке виконується для всіх $m \ge 1$. Індукційний крок: припустимо, що дане співвідношення справджується для всіх $n \le k$ $(k \ge 2)$, і розглянемо випадок $n = k + 1$. Послідовно матимемо $F_{k+1+m} = F_{k+m} + F_{k+m-1} = (F_{k-1}F_m + F_k F_{m+1}) + (F_{k-2}F_m + F_{k-1}F_{m+1}) = (F_{k-1} + F_{k-2})F_m + (F_k + F_{k-1})F_{m+1} = F_k F_m + F_{k+1}F_{m+1}$. (б) Використати метод математичної індукції за k і результат попередньої задачі. (в) Нехай d — найбільший спільний дільник чисел F_n і F_{n+1}. Зі співвідношення $F_{n+1} = F_n + F_{n-1}$ матимемо, що d — дільник F_{n-1}. Подібним чином встановимо, що d — дільник усіх попередніх чисел $F_{n-2}, F_{n-3}, \ldots, F_2, F_1$. Оскільки $F_2 = F_1 = 1$, то $d = 1$. (г) Замінимо в сумі перший доданок F_1 на F_2. Послідовно згортатимемо ліву частину, використовуючи такі рівності: $F_2 + F_3 = F_4$, $F_4 + F_5 = F_6$, …, $F_{2n} + F_{2n+1} = F_{2n+2}$. (д) Замінити перший доданок 1 на F_1. Див. (г). (е) Для перетворення лівої частини використати такі рівності: $F_n + F_{n+1} = F_{n+2}$, $F_{n+2} + F_{n+3} = F_{n+4}$, $F_{n+5} = F_{n+4} + F_{n+3}$. (є) Застосувати метод математичної індукції за n. Використати тотожність $F_{n+1}F_{n-1} - F_n^2 = (F_n + F_{n-1})F_{n-1} - F_n^2 = F_n(F_{n-1} - F_n) + F_{n-1}^2 = F_{n-1}^2 - F_n F_{n-2}$. (ж) Застосувати метод математичної індукції за n і використати біномні тотожності.

5.7. 1. 1) $(1-t)^{-1}$; 2) $(1-ct)^{-1}$; 3) $ct/(1-ct)^2$; 4) $(1-t)^{-2}$; 5) $ct/(1-t)^2$; 6) $t/(1-t)^2 + c/(1-t)$; 7) $t(1+t)/(1-t)^3$; 8) $tG'(t)$, де $G(t)$ — генератриса послідовності $b_k = k^{n-1}$; $k = 0, 1, 2, \ldots$; 9) $-\ln(1-t)/t$; 10) e^t; 11) e^{ct};

12) $(1-t)^{-(n+1)}$; **2.** 1) $(F(t)-a_0)/t$; 2) $(F(t)-(a_0+a_1t))/t^2$; 3) $\left(F(t)-\sum_{m=0}^{n-1} a_m t^m\right)/t^n$; 4) $F(t)+(1-t)^{-1}$; 5) $F(t)+c(1-t)^{-1}$; 6) $cF(t)$; 7) $cF(t)+d(1-t)^{-1}$; 8) $F(t)-\sum_{m=0}^{n-1} a_m t^m$; 9) $t^nF(t)$. 10) $F(t)(1-t)$; 11) $(F(t)(1-t)-a_0)/t$; 12) $F(t)(1-t)^{-1}$; 13) $tF'(t)$; 14) $tF'(t)+t^2F''(t)$; 15) $(\int_0^t F(x)dx)/t$. **3.** Послідовність β_n є результатом множення послідовностей β_{n-1} і (1, 1, ..., 1, ...). Її генератриса дорівнює $F(t)(1-t)^{-n}$, $n = 0, 1, 2, \dots$. **4.** 1) $a_0 = c$, $a_k = 0$ для $k > 0$; 2) $a_n = 1$ і $a_k = 0$ для $k \ne n$; 3) $a_k = 1$; 4) $a_k = k+1$; 5) $a_k = C(n+k-1, k)$; 6) $a_k = c^k$; 7) $a_k = kc^k$; 8) $a_k = ck$; 9) $a_k = k^2$; 10) $a_k = c^2k^2$; 11) $a_k = 0$, $k = 0, 1, \dots, n-1$; $a_k = C(k, n)$ для $k \ge n$; 12) $a_k = 0$, $k = 0, 1, \dots, n-1$; $a_k = C(k-1, n-1)$ для $k \ge n$; 13) $a_k = (-1)^k(1/2)_k/k!$; 14) $a_0 = 1$ і $a_k = (-4)^k(1/2)_k/k! = -2C(2k-2, k-1)/k$ для $k > 0$; 15) $a_k = C(n, k)$, $k = 0, 1, 2, \dots, n$, й $a_k = 0$ для $k > n$; 16) $a_k = C(n, k)c^kd^{n-k}$, $k = 0, 1, 2, \dots, n$, й $a_k = 0$ для $k > n$; 17) $a_k = 0$ для $k < n$, $k > 2n$ і $a_k = (-1)^{k-n}C(n, k-n)$ для $k = n, n+1, \dots, 2n$; 18) $a_k = 0$ для $k = 2m+1$ і $a_k = (1/2)^m(-n)_k/m!$ для $k = 2m$, $m = 0, 1, 2, \dots$; 19) $a_0 = 0$ і $a_k = (-1)^{k-1}/k$ для $k > 0$; 20) $a_0 = 0$ й $a_k = 1/k$ для $k > 0$; 21) $a_0 = 0$ й $a_k = c^k/k$ для $k > 0$; 22) $a_k = 0$ для $k = 2m$ і $a_k = c^k/k$ для $k = 2m+1$, $m = 0, 1, 2, \dots$; 23) $a_{2k} = 0$ і $a_{2k+1} = (-1)^k/(2k+1)$; 24) $a_{2k} = 0$; $a_1 = 1$ і $a_{2k+1} = (2k-1)!!/((2k)!!(2k+1))$ для $k > 0$. **5.** $F(t) = (1-t^n)^2(1-t)^{-2} = (1-2t^n+t^{2n})(1+2t+3t^2+\dots+kt^{k-1}+\dots)$. Шуканий коефіцієнт a_k при t^k дорівнюватиме $k+1$ для $k = 0, 1, 2, \dots, n-1$ і дорівнюватиме $2n-k-1$ для $k = n, n+1, \dots, 2n-2$; $a_k = 0$ для $k > 2n-2$. **6.** (а) $(1+t)^n = C(n, 0) + C(n, 1)t + C(n, 2)t^2 + \dots + C(n, n)t^n = \sum_{k=0}^{n} C(n, k)t^k$, $(1-t^2)^n = C(n, 0) - C(n, 1)t^2 + C(n, 2)t^4 - \dots + (-1)^nC(n, n)t^{2n}$, $(1-t)^{-n} = C(n-1, 0) + C(n, 1)t + C(n+1, 2)t^2 + \dots + C(n+k-1, k)t^k + \dots$. Перемножимо останні два розклади та визначимо коефіцієнт при t^k. Він дорівнюватиме $\sum_{l=0}^{[k/2]}(-1)^lC(n+k-2l-1, k-2l)C(n, l)$. Звідси випливає відповідна тотожність, бо $C(n+k-2l-1, k-2l) = C(n+k-2l-1, n-1)$. **7.** (а) Розглянемо тотожність $(1+t)^n(1-t)^n = (1-t^2)^n$ і визначимо коефіцієнт при t^n в обох її частинах. У лівій він дорівнюватиме $\sum_{k=0}^{n} C(n, k)(-1)^kC(n, n-k) = \sum_{k=0}^{n}(-1)^k(C(n, k))^2$, а в правій — $(-1)^{n/2}C(n, n/2)$,

якщо n — парне число, і 0, якщо n непарне. (б) Порівняти коефіцієнти при t^{n-1} у тотожності $(1+t)^n(1+t)^{-m-2} = (1+t)^{n-m-2}$. (в) $((1+t)^n + (1-t)^n)^2 = (1+t)^{2n} + 2(1-t^2)^n + (1-t)^{2n}$. (г) $((1+t)^n + (1-t)^n)((1+t)^n - (1-t)^n) = (1+t)^{2n} - (1-t)^{2n}$. **8.** (а) $F(t) = (f_0 + (f_1 + bf_0)t)/(1+bt+ct^2)$. (б) Якщо $1+bt+ct^2 = (1-a_1t)(1-a_2t)$ і $a_1 \ne a_2$, то $g(n) = D_1a_1^n - D_2a_2^n$, $D_i = (f_1 + (b+a_i)f_0)/(a_1 - a_2)$, $i = 1, 2$. У разі, коли $a_1 = a_2 = a$, то $g(n) = (f_0 + (f_1 - af_0)n/a)a^n$. **9.** (а) Розглянемо послідовність $\beta = (b_k)_{k=0,\infty} = \alpha \circ \alpha$. З означення операції множення послідовностей і заданого рекурентного співвідношення випливає, що $b_k = a_{k+1}$, $k = 0, 1, 2, \dots$. Отже, $\beta = \Delta_1\alpha$. Нехай $F(t)$ і $G(t)$ — генератриси послідовностей α та β відповідно. Тоді $G(t) = F(t)F(t) = F^2(t)$ і $G(t) = (F(t) - a_0)/t$ (див. (5.38)). Звідси $F^2(t) = (F(t)-1)/t$, тобто $F(t) = (1-(1-4t)^{1/2})/2t$. (б) Використати розклад функції $(1-4t)^{1/2}$ (див. 5.7.4(14)). $g(n) = C(2n, n)/(n+1)$. **10.** Позначимо $B(t) = (1+t)^n$. (а) Генератрисою послідовності $(1C(n, 1), 2C(n, 2), \dots, nC(n, n))$ є функцією $F(t) = B'(t) = n(1+t)^{n-1}$, тому шукана сума дорівнює $F(1) = n2^{n-1}$. (б) $F(t) = (tB(t))'$, $F(1) = (n+2)\cdot 2^{n-1}$. (в) $F(t) = ((B(t)-1)/t)'$, $F(1) = (n-2)2^{n-1}+1$. (г) $F(t) = 2B'(t) + B(t)$, $F(1) = (n+1)2^n$. (д) $F(t) = (t(1-t)^n)'$, $F(1) = 0$. (е) $F(t) = 4B'(t) - B(t) + 1$, $F(1) = (2n-1)2^n + 1$. (є) $F(t) = \int_0^t B(z)\mathrm{d}z$ — генератриса послідовності $(0, C(n, 0)/1, C(n, 1)/2, \dots, C(n, n)/(n+1))$, $F(1) = (2^{n+1}-1)/(n+1)$. (ж) $F(t) = \int_0^t (1-z)^n \mathrm{d}z$, $F(1) = 1/(n+1)$. **11.** Див. приклад 5.14. **12.** Нехай $F(t)$ — генератриса даної послідовності. Помноживши обидві частини заданого рекурентного співвідношення на t^n ($n = 0, 1, 2, \dots$) і додавши отримані рівності, матимемо $(F(t) - f_0 - f_1t)/t^2 + 2(F(t) - f_0)/t - 8F(t) = (1-2t)^{-1}$, тобто $F(t) = = ((1 - 4f_0 - 2f_1)t^2 + f_1t + f_0)/((1-2t)^2(1+4t)) = = (-t^2+2t)/(3(1-2t)^2(1+4t))$, або $F(t) = ((1-2t)^{-2} - (1+4t)^{-1})/12$. Розвинувши функцію $F(t)$ в ряд, отримаємо розв'язок рекурентного співвідношення: $g(n) = ((n+1)2^n - (-4)^n)/12$. **13.** (а) Позначимо послідовно вершини $(k+2)$-кутника символами $A_1, A_2, \dots, A_{k+2}$. Розглянемо множину розбиттів, які містять трикутник $A_1A_{k+2}A_m$, де $2 \le m \le k+1$. Кількість таких розбиттів дорівнює добутку числа розбиттів m-кутника $A_1A_2\dots A_m$ і числа розбиттів $(k-m+3)$-кутника $A_mA_{m+1}\dots A_{k+2}$. Перше з цих чисел дорівнює a_{m-2}, а друге — a_{k-m+1}. Змінюючи положення точки A_m, тобто надаючи m усіх можливих значень 2, 3, …, $k+1$, дістанемо таке рекурентне співвідношення: $a_k = a_0a_{k-1} + a_1a_{k-2} + a_2a_{k-3} + \ldots + a_{k-1}a_0$, $k = 1, 2, \ldots$ (для зручності покладемо $a_0 = 1$). Отже, $a_k = C(2k, k)/(k+1)$ (див. 5.7.9). (б), (в) Див. (а). (г) Нехай на площині проведено $k-1$ прямих. Проведемо ще одну пряму. Вона перетнеться з раніше проведеними прямими у $k-1$ точках, і розіб'ється ними на k частин, кожна з яких відповідатиме одній новій частині площини. Отже, $a_k = a_{k-1} + k$ і $a_0 = 1$. Звідси $a_k = k(k+1)/2 + 1$. **14.** (а) $\prod_{k=0}^{\infty}(1+t^k) - 1$. (б) $\prod_{k=1}^{\infty}(1-t^k)^{-1} - 1$. (в) $\prod_{k=1}^{\infty}(1+t^{2k-1}) - 1$. (г) $\prod_{k=1}^{\infty}(1-t^{2k-1})^{-1} - 1$. (д) $\prod_{k=1}^{m}(1-t^{k})^{-1}$. **15.** (а) $(1-t)^{-m}$, $a_n = C(m+n-1, n)$. (б) $t^m(1-t)^{-m}$, $a_n = 0$ для $n = 0, 1, \ldots, m-1$, і $a_n = C(n-1, m-1)$ для $n > m-1$.

6. Булеві функції

6.1. 1. (а) 2^{n-k}. (б) 2. (в) $C(n+1, 0) + C(n, 1) + C(n-1, 2) + \ldots + C(n-m+1, m)$, де $m = [(n+1)/2]$. Див. задачу 5.2.7. (г) $n+1$. (д) 2^{n-1}. Це число дорівнює $C(n, 0) + C(n, 2) + \ldots + C(n, 2k)$, $k = [n/2]$ (див. (5.8)). (е) 2^{n-1} (див. (д)). **2.** (а) (10101). (б) (1111111111). (в) (10011100010001). (г) (1110000). **3.** (а) (01101). (б) (0010100111). (в) (0010100000). (г) (1111111111). (д) (11…1) (m одиниць), бо $2^m - 1 = 2^{m-1} + 2^{m-2} + \ldots + 2^1 + 2^0$. (е) (10…0). (є) (100…01). (ж) (1011…1). (з) (11011…11). **4.** (а) $2^n - 2^k$. (б) 2^{k-1}. **5.** Нехай вектори $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ і $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ протилежні. Тоді $|a_i - b_i| = 1$ і $a_i + b_i = 1$ для всіх $i = 1, 2, \ldots, n$. Отже, $\nu(\boldsymbol{a}) + \nu(\boldsymbol{b}) = (a_12^{n-1} + a_22^{n-2} + \ldots + a_{n-1}2 + a_n) + (b_12^{n-1} + b_22^{n-2} + \ldots + b_{n-1}2 + b_n) = (a_1+b_1)2^{n-1} + (a_2+b_2)2^{n-2} + \ldots + (a_{n-1}+b_{n-1})2 + (a_n+b_n) = 2^{n-1} + 2^{n-2} + \ldots + 2 + 1 = 2^n - 1$. Припустимо, що для векторів $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ і $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ виконується рівність $\nu(\boldsymbol{a}) + \nu(\boldsymbol{b}) = 2^n - 1 = 2^{n-1} + 2^{n-2} + \ldots + 2 + 1$. Розписавши $\nu(\boldsymbol{a})$ та $\nu(\boldsymbol{b})$ за формулами, дістанемо $a_i + b_i = 1$, або $|a_i - b_i| = 1$, для $i = 1, 2, \ldots, n$. Отже, $\rho(\boldsymbol{a}, \boldsymbol{b}) = n$. **6.** Нехай вектори $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ і $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ — сусідні. Тоді існуватиме такий номер k, що $|a_k - b_k| = 1$ і $|a_i - b_i| = 0$ для всіх $i \neq k$. Отже, $|\nu(\boldsymbol{a}) - \nu(\boldsymbol{b})| = |a_k - b_k|2^{n-k} = 2^{n-k}$. Обернене твердження неправильне. Наприклад, якщо $\nu(\boldsymbol{a}) = 2^k$ і $\nu(\boldsymbol{b}) = 2^{k-1}$, то $|\nu(\boldsymbol{a}) - \nu(\boldsymbol{b})| = 2^{k-1}$, однак $\boldsymbol{a}$ і $\boldsymbol{b}$ не є сусідніми векторами, бо $\rho(\boldsymbol{a}, \boldsymbol{b}) = 2$. **7.** (а) $C(n, \|\boldsymbol{a}\|)$. (б) n. (в) 1.

(г) $C(n, k+1) + C(n, k+2) + \ldots + C(n, n)$. **8.** (а) $n2^{n-1}$. Для кожного з 2^n векторів множини B^n існує n сусідніх. Отже, є $n2^n$ упорядкованих пар сусідніх векторів, а кількість невпорядкованих удвічі менша. (б) 2^n. Для кожного з 2^n векторів існує єдиний протилежний йому. (в) $C(n, k)2^n$. **9.** Відношення рівності ваг рефлексивне, симетричне і транзитивне, тому R — еквівалентність. Індекс відношення R дорівнює $n+1$. Якщо $\|a\| = k$, то кількість елементів у класі $[a]_R$ — $C(n, k)$. **10.** $C(2^{n+1} - 1, 2^n - 1)$. Див. задачу 5.2.8. **11.** (а) x_2, x_3; $\boldsymbol{a}_g = (10)$. (б) x_3, x_4; $\boldsymbol{a}_g = (1001)$. **12.** Якщо $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ — фіктивні змінні функції f і $f(a_1, a_2, \ldots, a_n) = 1$, то f набуває значення 1 також на всіх $2^k - 1$ наборах, які відрізняються від $(a_1, a_2, \ldots, a_n)$ зміною значень координат з номерами $i_1, i_2, \ldots, i_k$. Отже, дане твердження можна підсилити: якщо функція f має не менше k фіктивних змінних, то $|N_f|$ кратне 2^k.

6.2. 1. (а) (00010111). (б) (11101000). (в) Наприклад, (01010101). (г) Наприклад, (01100110). **2.** Поставимо у відповідність членові A_i булеву змінну x_i, вважаючи, що $x_i = 1$ тоді й тільки тоді, коли A_i бере участь у засіданні комітету, $i = 1, 2, 3, 4$. Запишемо умови відвідування засідань за допомогою булевої функції: $f(x_1, x_2, x_3, x_4) = x_2\overline{x}_3x_4 \vee x_1\overline{x}_2x_3 \vee x_1x_4$. Якщо A_3 не бере участі в засіданні (тобто $x_3 = 0$), то для з'ясування того, чи A_2 обов'язково повинен бути на цьому засіданні, слід визначити, чи функція $f(x_1, x_2, 0, x_4)$ набуває значення 1 на якомусь наборі виду $(a_1, 0, 0, a_4)$. Оскільки $f(x_1, 0, 0, x_4) = x_1x_4$, то $f(1, 0, 0, 1) = 1$, тому засідання комітету може відбутися без участі як A_3, так і A_2. **3.** Поставимо у відповідність повідомленням ***A***, ***B*** й ***C*** булеві змінні x, y і z відповідно. Вважаємо, що булева змінна дорівнює 1 тоді й тільки тоді, коли передано відповідне повідомлення. Подіям 1), 2) і 3) відповідатимуть вирази $(\overline{x \wedge y})$, $x \sim (y \wedge z)$ та $x \vee z$, а загальну ситуацію описує булева функція $f(x, y, z) = (\overline{x \wedge y}) \wedge (x \sim (y \wedge z)) \wedge (x \vee z)$. Для відповіді на поставлене питання слід з'ясувати, чи формула $f(x, y, z) \to (\overline{y} \wedge z)$ тотожно дорівнює 1. Відповідь позитивна. **4.** (а) (1011). (б) (10001111). **5.** (а), (б), (в) 2^{n-1}. **6.** (а) $2^m - 1$, де $m = 2^n$. (б) $2^m - k$, де $m = 2^n$. (в) 2^p, де $p = 2^{n-1}$. **7.** 2^t, де $t = 3 \cdot 2^{n-2}$. Функція f, що задовольняє умови задачі, може набувати довільних значень на $2 \cdot 2^{n-2}$ наборах, у яких $(x_i, x_j) = (0, 0)$ або $(x_i, x_j) = (1, 1)$. Окрім того, значення функції f на наборі, в якому $(x_i, x_j) = (0, 1)$ (число таких наборів — 2^{n-2}), має збігатися зі значенням f на наборі, де $(x_i, x_j) = (1, 0)$. Отже, для задання функції f за допомогою її стовпчика (вектора) значень існує $3 \cdot 2^{n-2}$ позицій вибору з двох можливих варіантів: 0 або 1. **8.** $|N_f|$ — кількість наборів значень, на яких функція f дорівнює 1, а $\|\boldsymbol{a}_f\|$ — кількість одиниць у векторі значень функції f.

6.3. 1. (а) Формула. (б) Не формула, але за допомогою додавання дужок може бути перетворена у формулу. **2.** (а) $f_1(x, f_3(x, y, z))$, $f_3(f_1(x, y), z, f_1(x, z))$, $f_3(x, y, z)$, $f_1(x, y)$, $f_1(x, z)$. (в) $f_1(x, (x \to y))$, $f_2(x, f_1(y \oplus z, x)) \to z$, $x \to y$, $f_2(x, f_1(y \oplus z, x))$, $f_2(x, f_1(y \oplus z, x))$, $f_1(y \oplus z, x)$, $y \oplus z$. **3.** Застосувати індукцію за означенням формули (суперпозиції). **5, 6.** Обидва твердження є наслідками індуктивного правила побудови формули. **7.** (а) (10100000). (б) (10011000). (в) (00100010). **8.** (а) (1101). (б) (1100111100001101). **9.** Розв'язання задач такого типу потребує систематичного перебору формул, утворених у результаті суперпозицій функцій із множини D. Такої систематичності можна досягти за допомогою, наприклад, упорядкування отримуваних формул за складністю або за кількістю кроків їх побудови. (а) $\{1, x, y, x \to y, y \to x, x \wedge y, x \vee y, x \sim y\}$. (б) $\{0, 1, x, y, x \oplus 1, y \oplus 1, x \oplus y, x \oplus y \oplus 1\}$. (в) $P_2(2)$.

6.4. 1. (а)–(д) рівносильні. **5.** (б) Довести, що, наприклад, формули $(x \to y) \to z$ і $x \to (y \to z)$ нерівносильні. (е) Довести, що, наприклад, формули $x \vee (\overline{y \oplus z})$ і $(x \vee y) \oplus (x \vee z)$ нерівносильні. **7.** (а) $\overline{x \vee y} = \bar{x} \wedge \bar{y}$. (б) $\overline{x \to y} = x \wedge \bar{y}$. (в) $x \vee y = \neg x \to y$. (г) $x \oplus y = \neg(\neg(x \vee y) \vee \neg(\neg x \vee \neg y))$. (д) $x \oplus y = (\bar{x} \to y) \to \overline{x \to y}$, див. (в) і (г). (е) $x \vee y = (x \mid x) \mid (y \mid y)$. (є) Див. (г) і (е). (ж) $x \to y = x \mid (y \mid y)$. (з) $x \wedge y = (x \downarrow x) \downarrow (y \downarrow y)$. (и) $x \to y = ((x \downarrow x) \downarrow y) \downarrow ((x \downarrow x) \downarrow y)$. (і) $x \mid y = ((x \downarrow x) \downarrow (y \downarrow y)) \downarrow ((x \downarrow x) \downarrow (y \downarrow y))$. (ї) $x \sim y = (x \to y) \wedge (y \to x)$. **8.** (а) Від супротивного. Припустімо, що існує формула $F(x_1, x_2, \ldots, x_n)$ над H, рівносильна $\bar{x}_1$, тобто виконується тотожність $\bar{x}_1 = F(x_1, x_2, \ldots, x_n)$. Підставимо в цю тотожність замість змінних набір значень $(1, 1, \ldots, 1)$. Оскільки результати всіх операцій з H дорівнюють 1, коли їх операнди набувають значення 1, то $F(1, 1, \ldots, 1) = 1$. У той же час у лівій частині тотожності дістанемо $\bar{1} = 0$. Отримали суперечність, яка спростовує припущення про існування формули над H, що задає функцію заперечення. (б), (в) Див. (а). Розглянути набір значень $(0, 0, \ldots, 0)$. (г) Див. (а). **9.** (а) $\{0\}$. (б) $\{(0, 0, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 1)\}$. **10.** $\{(1, 1, 1)\}$. **11.** (а) $2^n - 1$. Рівняння задовольняють усі набори з B^n, окрім $(0, 0, \ldots, 0)$.

(б) 2^{n-1}. Рівняння задовольняють двійкові вектори, вага яких парна (див. 6.1.1(д)). (в) 2. (г) $2^{n-1} + 1$. Рівняння задовольняють усі двійкові вектори, вага яких непарна, а також вектор $(0, 0, \ldots, 0)$ (див. 6.1.1(е)). (д) $2^{n-1} - 1$, якщо n непарне; 2^{n-1}, якщо n парне. (е) $2^{n-1} + 1$, якщо n непарне; $2^{n-2} + 1$, якщо n парне. **12.** (а) 2^{n-1}. Будь-який розв'язок другого рівняння задовольняє перше рівняння. (б) Один розв'язок $(1, 1, \ldots, 1)$, якщо n парне, і жодного розв'язку, якщо n непарне. Розв'язками першого рівняння є набори $(0, 0, \ldots, 0)$ і $(1, 1, \ldots, 1)$. (в) Два розв'язки $(0, 0, \ldots, 0)$ і $(1, 1, \ldots, 1)$, коли n непарне, і один розв'язок $(0, 0, \ldots, 0)$, коли n парне (див. (б)). (г) $2^{n-1} - 2$ для непарного n і $2^{n-1} - 1$ для парного. Перше рівняння задовольняють усі двійкові вектори, окрім $(0, 0, \ldots, 0)$ і $(1, 1, \ldots, 1)$. Друге рівняння задовольняють двійкові вектори, кількість одиниць у яких парна, причому вектор $(1, 1, \ldots 1)$ буде розв'язком лише для непарного n. **13.** (а) $\{0\}$. (б) $\{(0, 0)\}$. **14.** (а) Один розв'язок $(1, 1, \ldots, 1)$, якщо n парне, жодного розв'язку — в іншому разі. (б) K, якщо n парне; $K + 1$ — в іншому разі, де $K = 2^{n-2} + 2^{n-4} + \ldots + 2^{n-2k} = 2^{n-2k}(2^{2k} - 1)/3$, $k = [n/2]$. Розв'язками нерівності будуть двійкові вектори $(a_1, a_2, \ldots, a_n)$, в яких $a_{2k} = 0$ і $a_1 = a_2 = \ldots = a_{2k-1} = 1$ для $k = 1, 2, \ldots, [n/2]$, а також вектор $(1, 1, \ldots, 1)$, якщо n непарне. **15.** Нехай задано булеві функції f і g. Неважко переконатися, що $f \le g$ тоді й лише тоді, коли $\boldsymbol{a}_f \le \boldsymbol{a}_g$ (див. приклад 1.17(4)). **16.** Пара нерівностей $\boldsymbol{a}_f \le \boldsymbol{a}_g$ і $\boldsymbol{a}_g \le \boldsymbol{a}_f$ рівносильна рівності $\boldsymbol{a}_f = \boldsymbol{a}_g$ чи $f = g$ (див. попередню задачу).

6.5. 1. (а) $f(x_1, x_2, x_3) = x_1 f_1(x_2, x_3) \vee \bar{x}_1 f_2(x_2, x_3)$, де $\boldsymbol{a}_{f_1} = (0111)$, $\boldsymbol{a}_{f_2} = (0010)$. (б) $f(x_1, x_2, x_3, x_4) = x_1 x_3 f_{11}(x_2, x_4) \vee x_1 \bar{x}_3 f_{12}(x_2, x_4) \vee \bar{x}_1 x_3 f_{21}(x_2, x_4) \vee \bar{x}_1 \bar{x}_3 f_{22}(x_2, x_4)$, де $\boldsymbol{a}_{f_{11}} = (1011)$, $\boldsymbol{a}_{f_{12}} = (1000)$, $\boldsymbol{a}_{f_{21}} = (0110)$, $\boldsymbol{a}_{f_{22}} = (0110)$. **2.** (а) $x_1(x_2 \oplus x_3) \vee \bar{x}_1 x_2$. (б) $x_2 x_4 \vee x_2 \bar{x}_4 x_1 x_3 \vee \bar{x}_2 x_4 \vee \bar{x}_2 \bar{x}_4 x_1$. **3.** Перевірити виконання обох співвідношень для наборів значень змінних, що мають вигляд $(0, a_2, a_3, \ldots, a_n)$ і $(1, a_2, a_3, \ldots, a_n)$. **4.** (а) $\bar{x}_1 x_2$. (б) $\bar{x}_1 \bar{x}_2 \bar{x}_3 \vee x_1 \bar{x}_2 \bar{x}_3 \vee x_1 x_2 x_3$. (в) $\bar{x}_1 \bar{x}_2 x_3 x_4 \vee \bar{x}_1 x_2 x_3 x_4 \vee x_1 \bar{x}_2 \bar{x}_3 x_4 \vee x_1 x_2 \bar{x}_3 x_4 \vee x_1 x_2 x_3 \bar{x}_4$. **5.** Побудувати спочатку таблицю істинності для формули A, відтак шукану ДДНФ. (а) $\bar{x} y \vee x \bar{y}$. (б) $\bar{x}\bar{y}\bar{z} \vee \bar{x}\bar{y}z \vee \bar{x}y\bar{z} \vee \bar{x}yz$. **6.** Для формули, що зображує булеву функцію f, яка не дорівнює тотожно 0, такою ДНФ є, наприклад, ДДНФ функції f, в іншому разі відповідна ДНФ має вигляд 0 (константу 0 називають елементарною кон'юнкцією рангу

0). **7.** Використати тотожності задачі 6.4.2 та основні тотожності (6.3). (а) $\overline{x}\,\overline{y}\,\overline{z} \vee xy\overline{z}$ (б) $\overline{x}\,\overline{z} \vee xyz$.

6.6. 1. Відношення R рівносильності формул означено через рівність відповідних функцій, яка має властивості рефлексивності, симетричності і транзитивності. **2, 3.** Застосувати індукцію за числом кроків побудови формули C. **4.** Шукану взаємно однозначну відповідність φ між Φ/R і P означимо так: для довільної формули F над H $\varphi([F]) = f$, де f — функція, яку реалізує F. Нехай σ — якась n-арна операція з H, $f_1, f_2, \ldots, f_n \in P$ і формули $F_1, F_2, \ldots, F_n$ зображають функції $f_1, f_2, \ldots, f_n$ відповідно. Тоді результатом h відповідної операції σ' для $f_1, f_2, \ldots, f_n$ вважаємо функцію з P, яку реалізує формула $\sigma(F_1, F_2, \ldots, F_n)$; записуємо $h = \sigma'(f_1, f_2, \ldots, f_n)$. **5.** Наприклад, умови (2.1) для відображення γ, будь-якої n-арної операції σ з H і довільних формул $F_1, F_2, \ldots, F_n$ над H матимуть вигляд $\gamma(\sigma(F_1, F_2, \ldots, F_n)) = \sigma'(\gamma(F_1), \gamma(F_2), \ldots, \gamma(F_n))$. Виконання цих рівностей грунтується на означеннях відображення γ і відповідних операцій σ' (див. попередню задачу). Отже, γ — гомоморфізм. Ізоморфізм відображення δ випливає з того, що δ — бієкція і для δ виконуються умови (2.1), бо для довільної n-арної операції σ фактор-алгебри D/R і довільних класів рівносильних формул $[F_1], [F_2], \ldots, [F_n]$ виконується рівність $\delta(\sigma([F_1], [F_2], \ldots, [F_n])) = \sigma'(\delta([F_1]), \delta([F_2]), \ldots, \delta([F_n]))$. Для обгрунтування останньої рівності слід використати співвідношення $\sigma([F_1], [F_2], \ldots, [F_n]) = [\sigma(F_1, F_2, \ldots, F_n)]$ і $\delta([F]) = \gamma(F)$ та вищенаведені міркування.

6.7. 1. (а) $\overline{x}\,\overline{y} \vee \overline{x}y \vee xy$. (б) $\overline{x}\,\overline{y}z \vee \overline{x}y\overline{z} \vee \overline{x}yz \vee x\overline{y}\,\overline{z} \vee xyz$. (в) $\overline{x}\,\overline{y}\,\overline{z} \vee \overline{x}\,\overline{y}z \vee \vee \overline{x}y\overline{z} \vee \overline{x}yz \vee x\overline{y}z \vee xy\overline{z} \vee xyz$. **3.** (а) Утворити диз'юнкцію всіх елементарних кон'юнкцій рангу n, які не ввішли до D_1. (б) Визначити множину всіх елементарних кон'юнкцій, що входять або до D_1, або до D_2 і утворити диз'юнкцію цих елементарних кон'юнкцій. (в) Утворити диз'юнкцію елементарних кон'юнкцій, які одночасно входять до D_1 і D_2. (г) Перетворити формулу $D_1 \oplus D_2$ до рівносильної формули $(\mathrm{D}_1 \wedge D_2) \vee (D_1 \wedge \mathrm{D}_2)$ та використати результати попередніх задач. (д) $D_1 \to D_2 = \mathrm{D}_1 \vee D_2$, далі див. (а) і (б). (е) $D_1 \sim D_2 = (\mathrm{D}_1 \wedge \mathrm{D}_2) \vee (D_1 \wedge D_2)$, далі див. (а), (б) і (в). **4.** (а) $C = \mathrm{D}_1 \wedge D_2$, далі див. 6.7.3(а, в). (б) $C = \mathrm{D}_1 \vee \mathrm{D}_2$, далі див. 6.7.3(а, б).

5. На будь-якому наборі значень змінних не більше ніж одна елементарна кон'юнкція в ДДНФ набуває значення 1, тому після заміни $\vee$ на $\oplus$ дістаємо рівносильну формулу. Для довільної ДНФ зазначена властивість не виконується (розглянути, наприклад, $xy \vee yz$). **6.** (а) $(\overline{x}_1 \vee x_2) \wedge (\overline{x}_1 \vee \overline{x}_2)$. (б) $(x_1 \vee x_2 \vee x_3) \wedge \wedge (x_1 \vee \overline{x}_2 \vee x_3) \wedge (\overline{x}_1 \vee x_2 \vee x_3) \wedge \wedge(\overline{x}_1 \vee x_2 \vee \overline{x}_3)$. (в) $(x_1 \vee x_2 \vee \overline{x}_3 \vee x_4) \wedge (x_1 \vee \overline{x}_2 \vee x_3 \vee x_4) \wedge (\overline{x}_1 \vee x_2 \vee x_3 \vee x_4) \wedge \wedge (\overline{x}_1 \vee x_2 \vee \overline{x}_3 \vee x_4) \wedge (\overline{x}_1 \vee x_2 \vee \overline{x}_3 \vee \overline{x}_4) \wedge (\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3 \vee \overline{x}_4)$. **7.** $(x \vee y \vee z) \wedge (x \vee y \vee \overline{z}) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee y \vee z)$. **8.** Після виконання першого кроку дістанемо ДКНФ формули $\overline{A}$, а після виконання другого — ДДНФ формули $\overline{\overline{A}} = A$. **9.** 1) Виписати диз'юнкцію всіх елементарних кон'юнкцій, що не ввійшли до A. 2) Замінити $\wedge$ на $\vee$, $\vee$ — на $\wedge$, x_i — на $\overline{x}_i$, $\overline{x}_i$ — на x_i. **10.** (а) ДДНФ формули $\overline{A}$ — це диз'юнкція всіх елементарних кон'юнкцій, що не ввійшли до ДДНФ формули A. ДКНФ формули $\overline{A}$ отримаємо, якщо в ДДНФ формули A замінимо $\wedge$ на $\vee$, $\vee$ — на $\wedge$, x_i — на $\overline{x}_i$, $\overline{x}_i$ — на x_i. (б) Для ДДНФ див. 6.7.3(б). Отримавши ДДНФ формули $(A \vee B)$, будуємо ДКНФ за процедурою з 6.7.9. (в) Для ДДНФ див. 6.7.3(в), для ДКНФ — 6.7.9. (г) $A \to B = \overline{A} \vee B$, далі див. (а) та (б).

6.8. 1. (а) (1100). (б) (01111111). (в) (0111110011101010). **2.** (а), (б), (г) — так. **3.** За правилами побудови вектора значень $\boldsymbol{a}_f$, якщо a_j — значення функції f на наборі $(b_1, b_2, \ldots, b_n)$, то $j = \nu((b_1, b_2, \ldots, b_n)) = b_1 2^{n-1} + b_2 2^{n-2} + \ldots + b_{n-1} 2 + b_n$. Нехай $(c_1, c_2, \ldots, c_n)$ — довільний двійковий набір, і $\nu((c_1, c_2, \ldots, c_n)) = l$. Тоді $f^*(c_1, c_2, \ldots, c_n) = \overline{f(\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_n)} = \overline{a}_m$ і $m = \nu((\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_n)) = (1 - c_1) \cdot 2^{n-1} + (1 - c_2) 2^{n-2} + \ldots + (1 - c_{n-1}) 2 + (1 - c_n) (2^n - 1) - (c_1 2^{n-1} + c_2 2^{n-2} + \ldots + c_{n-1} 2 + c_n) = (2^n - 1) - \nu((c_1, c_2, \ldots, c_n)) = (2^n - 1) - l$. Отже, l-й елемент (нумерація починається з 0) вектора значень функції f^* збігається із запереченням елемента, розміщеного на l-й позиції від правого кінця у векторі $\boldsymbol{a}_f$. **4.** (а) Нехай $a_f = (a_0, a_1, \ldots, a_k)$, $k = 2^n - 1$. Оскільки $|N_f| = \|\boldsymbol{a}_f\|$ (див. 6.2.8), то з результату попередньої задачі випливає, що $|N_f| + |N_g| = (a_0 + a_1 + \ldots + a_k) + (\overline{a}_k + \overline{a}_{k-1} + \ldots + \overline{a}_0) = (a_0 + \overline{a}_0) + (a_1 + \overline{a}_1) + \ldots + (a_k + \overline{a}_k) = 2^n$, бо $b + \overline{b} = 1$ для $b \in B$. (б) Оскільки $|N_f| + |N_g| = 2^n$ (див. (а)), то $|N_f| - |N_g| = 2|N_f| - 2^n$. (в) Оскільки $|b - c| = |c - \overline{b}|$ для $c, b \in B$, то $\rho(\boldsymbol{a}_f, \boldsymbol{a}_g) = |a_0 - \overline{a}_k| + |a_1 - \overline{a}_{k-1}| + \ldots + |a_k - \overline{a}_0| = 2(|a_0 - \overline{a}_k| + |a_1 - \overline{a}_{k-1}| + \ldots + |a_m - \overline{a}_{k-m}|)$, де $m = 2^{n-1} - 1$. **5.** Порівняти відповідні таблиці істинності. **6.** Справедливість твердження випливає з того, що коли $f = g$, то $f^* = g^*$, і навпаки.

6.9. 1. (а) Див. 6.4.8(г). (б) Можна довести (наприклад, індукцією за кількістю операцій $\rightarrow$), що будь-яка формула над $\{\rightarrow\}$ набуває значення 1 не менше ніж на половині наборів значень своїх операндів. Формула A не має такої властивості. (в) Див. (б). (г) Можна довести (наприклад, індукцією за числом операцій $\vee$ і $\rightarrow$), що функція $g(x_1, x_2, \ldots, x_n)$, задана довільною формулою над $\{\vee, \rightarrow\}$, має властивість: існує таке k, що $g(x_1, x_2, \ldots, x_k, \ldots, x_n) \geq x_k$. Функція f не задовольняє цю умову. **2.** (а) $x \vee y = \bar{x} \rightarrow y$, $x \wedge y = \overline{x \rightarrow \bar{y}}$ (б) $0 = (x \wedge x) \oplus x$, $x \wedge y = (x \wedge y) \oplus 0$, $\bar{x} = (x \sim x) \oplus x$. (в) $\bar{x} = x \oplus x \oplus x$, $x \vee y = \bar{x} \rightarrow y$. (г) $\bar{x} = x \rightarrow 0$, $x \vee y = \bar{x} \rightarrow y$. (д) $\bar{x} = x \oplus 1$. (е) $0 = x \oplus x$, далі див. (г). **3.** Див. приклад 6.1(4) або задачу 6.4.8(а).

6.10. 1. 1, якщо $k = 0$; $K = 2^S(2^{C(n,\,k)} - 1)$ для $k > 0$, де $S = C(n, 0) + C(n, 1) + \ldots + C(n, k - 1)$. **2.** 1, якщо $k = 0$, і $K/2$ для $k > 0$ (див. попередню задачу). **3.** $C(2^n, m)$. **4.** Застосувати стандартний метод перевірки рівносильності формул. **5.** (а) $xy \oplus x \oplus 1$. (б) $xyz \oplus xy \oplus z$. (в) $xyzw \oplus xyw \oplus yzw \oplus xzw \oplus yw \oplus w \oplus 1$. **6.** (а) $xy \oplus x \oplus y$. (б) $xyz \oplus xy \oplus xz \oplus z$. (в) $xyw \oplus yzw \oplus xw \oplus zw \oplus 1$. **7.** Необхідність: якби істотна змінна x_i не входила явно в поліном Жегалкіна функції f, то зміна значення x_i не впливала б на значення функції f, і це суперечило б істотності змінної x_i. Достатність. Нехай x_i входить у поліном Жегалкіна $P(x_1, x_2, \ldots, x_n)$ функції f. Подамо поліном P у вигляді $P(x_1, x_2, \ldots, x_n) = x_i Q(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \oplus R(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. Оскільки x_i явно входить у P, то поліном Q не дорівнює тотожно 0. Тому існує набір значень $(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$, на якому Q дорівнює 1. Підставимо цей набір у вищенаведений розклад, дістанемо $g(x_i) = P(a_1, a_2, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n) = x_i \oplus b$, де $b = R(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$. Отже, $(a_1, a_2, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) = g(0) = b \neq 1 \oplus b = g(1) = f(a_1, a_2, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n)$, тому x_i — істотна змінна функції f.

6.11. 1. Див. зауваження до розв'язання задач 6.3.9. (а) $\{x, \bar{x}, y, \bar{y}\}$. (б) $\{0, x, y, x \oplus y\}$. (в) $\{0, 1, x, y, \bar{x}, \bar{y}\}$. (г) $\{1, x, y, x \rightarrow y, y \rightarrow x, x \vee y\}$. (д) $\{x, y\}$. (е) $\{0, 1, x, y, \bar{x}, \bar{y}, x \sim y, x \oplus y\}$. **2.** (а) $f(x) = x \rightarrow 0$. (б) $f(x, y) = ((x \downarrow x) \downarrow (y \downarrow y)) \downarrow (x \downarrow y)$. (в) $f(x, y) = (x \rightarrow y) \rightarrow y$. **3.** (а), (д) — замкнені. Наприклад, для функцій φ_2, φ_4 і φ_8, що належать множинам (б) та (є), функція $\varphi_8(\varphi_2(x, y), \varphi_4(x, y))$ не входить у ці множини (див. табл. 6.3). Для (в) розглянути функцію $\varphi_6(\varphi_6(x, y), \varphi_9(x, y))$, а для (е) — $\varphi_1(\varphi_2(x, y), y)$. У результаті суперпозиції поліномів із множини (г) можна отримати поліном степеня більшого, ніж 2. **4.** (а) Застосувати метод, подібний до використаного в доведенні лем 6.1–6.4. (б) Розглянути, наприклад, об'єднання $T_0 \cup T_1$ і суперпозицію $(x \vee y) \rightarrow (x \oplus y)$. **5.** (а) Включення $[K] \subseteq [[K]]$ виконується за означенням замикання. Нехай $f \in [[K]]$. Тоді f є суперпозицією функцій $g_1, g_2, \ldots, g_k$ з $[K]$. У свою чергу, кожна з функцій g_i є суперпозицією якихось функцій $h_{i1}, h_{i2}, \ldots, h_{im_i}$ з K, $i = 1, 2, \ldots, k$. Звідси внаслідок індуктивного означення суперпозиції можна твердити, що f є суперпозицією функцій h_{ij}, $j = 1, 2, \ldots, m_i$, $i = 1, 2, \ldots, k$, які належать K. Отже, $f \in [K]$. (б) Нехай $f \in [K_1]$. Тоді f є суперпозицією функцій $g_1, g_2, \ldots, g_k \in K_1$. З умови випливає, що всі ці функції належать також K_2. Тому $f \in [K_2]$. (в), (г) Див. (б). **6.** Класу $T_0 \cup T_1$ належать (а), (б), (в); класу $T_1 \setminus T_0$ — не належить жодна з функцій; класу $T_0 \setminus T_1$ — належить (а), (в); класу $T_0 \cap T_1$ — (б). **7.** Обчислити значення формули A на наборах (0, 0, 0) і (1, 1, 1). **8.** (а) $f \in T_0$ для будь-якого n і $f \in T_1$ для непарного n. Отже, $f \in T_0 \cup T_1$ і $f \notin T_1 \setminus T_0$ для всіх n; $f \in T_0 \setminus T_1$ для парних n і $f \in T_0 \cap T_1$ для непарних n. (б) Див. (а). (в) $f \in T_0$ для будь-якого n і $f \in T_1$ для $n = 4k + 2$ або $n = 4k + 3$, $k \in N$, бо кількість доданків у формулі A дорівнює $C(n, 2) = n(n-1)/2$. Далі див. (а). (г) $f \in T_0$ для всіх n і $f \in T_1$ для $n = 4k + 3$, $k \in N$, бо число доданків в A дорівнює $C(n, 3) = n(n-1)(n-2)/6$. Далі див. (а). **9.** (а) Нехай $f(x_1, x_2, \ldots, x_n) \in T_0^*$. Це означає, що існує така функція $g(x_1, x_2, \ldots, x_n) \in T_0$, що $g^* = f$. Звідси $f(1, 1, \ldots, 1) = g^*(1, 1, \ldots, 1) = g(0, 0, \ldots, 0) = 1$, тобто $f \in T_1$. Навпаки, припустімо, що $f(x_1, x_2, \ldots, x_n) \in T_1$. Тоді $f^*(0, 0, \ldots, 0) = \bar{f}(1, 1, \ldots, 1) = 0$, тобто $f^* \in T_0$. Звідси випливає $f \in T_0^*$, бо $(f^*)^* = f$. (б) Див. (а). **10.** Позначимо $K = 2^n$. (а), (б) 2^{K-1}. Наприклад, будь-яка функція $f(x_1, x_2, \ldots, x_n)$ із класу T_0 може набувати довільних значень (0 або 1) на $2^n - 1$ наборах значень своїх змінних (окрім набору (0, 0, …, 0)). (в) 2^{K-2}. (г) $3 \cdot 2^{K-2}$. $|T_0 \cup T_1| = |T_0| + |T_1| - |T_0 \cap T_1|$ (див. (а)–(в)). (д) 2^{K-2}. $|T_0 \setminus T_1| = |T_0| - |T_0 \cap T_1|$ (див. (а), (в)). (е) 2^{K-2} (див. (д)). (є) 2^{K-1} (див.(д), (е)). (ж) 2^{K-1}. (з) $3 \cdot 2^{K-2}$ (див.(в)). (и) 2^{K-2} (див. (г)). **11.** (а), (б), (в) — нелінійні. **12.** Побудувати відповідну таблицю істинності функції f, відтак — поліном Жегалкіна. (а) Лінійна. (б) Нелінійна. **13.** Застосувати метод невизначених коефіцієнтів. (а) (1001), $x \oplus y \oplus 1$. (б) (10010110), $x \oplus y \oplus z \oplus 1$. (в) (10011001), $y \oplus z \oplus 1$. **14.** Застосуємо метод доведення від супротивного. Припустімо, що якась змінна x_i фіктивна. Тоді на сусідніх наборах, які відрізняються лише i-ю координатою, функція f набуватиме однакових значень, що суперечитиме умові. Припустімо, що f — нелінійна функція. Тоді шляхом підстановки замість $n-2$ її змінних певних булевих констант f можна перетворити у нелінійну бінарну булеву функцію $g(x_i, x_j) = x_i x_j \oplus a x_i \oplus b x_j \oplus c$ (див. лему 6.7). Використовуючи умову задачі, зі співвідношень $g(0, 0) = g(0, 1) = g(1, 0)$, $g(1, 1) = g(0, 1) = g(1, 0)$ дістанемо систему рівнянь: $c = \overline{b \oplus c} = \overline{a \oplus c}$, $1 \oplus a \oplus b \oplus c = \overline{b \oplus c} = \overline{a \oplus c}$. Ця система несумісна. Обернене твердження неправильне (розглянути, наприклад, функцію $f(x_1, x_2, x_3) = x_1 \oplus x_3$). **15.** Якщо $f(x_1, x_2, \dots, x_n)$ — лінійна функція, то її поліном Жегалкіна має вигляд $a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus a_{n+1}$. Оскільки f відмінна від константи, то принаймні один із коефіцієнтів a_k дорівнює 1, $k = 1, 2, \dots, n$. Функція f набуватиме протилежних значень на наборах, що відрізняються лише k-ю координатою. Отже, кількість наборів, на яких f набуває значення 1, дорівнює половині від усіх можливих, тобто $|N_f| = 2^n/2 = 2^{n-1}$. Обернене твердження неправильне (див., наприклад, 6.11.11(б, в)). **16.** $2C(n, k)$. Якщо $f(x_1, x_2, \dots, x_n)$ — лінійна функція, то її поліном Жегалкіна має вигляд $a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus a_{n+1}$. Функція f істотно залежить від x_i тоді й тільки тоді, коли $a_i = 1$ (див. 6.10.7). Отже, існує $C(n, k)$ способів обрати k істотних змінних серед n можливих і 2 способи обрати значення вільного члена a_{n+1}. **17.** Позначимо $K = 2^n$. (а) 2^{n+1}. Це число способів обрати значення (із двох можливих) для $n+1$ коефіцієнта $a_1, a_2, \dots, a_n, a_{n+1}$ відповідного полінома Жегалкіна. (б) 2^n. Вільний член a_{n+1} відповідного полінома Жегалкіна має дорівнювати 0. (в) 2^n. Серед коефіцієнтів відповідного полінома Жегалкіна кількість одиничних має бути непарною. Отже, шукане число дорівнює $C(n+1, 1) + C(n+1, 3) + \dots + C(n+1, 2k+1)$, де $k = [(n-1)/2]$ (див. біномні тотожності (5.8)). (г) $2^{K-1} + 2^n$. $|T_0 \cup L| = |T_0| + |L| - |T_0 \cap L|$ (див. 6.11.10(а) і пункти (а) і (б) цієї задачі). (д) 2^{n-1} (див. (б) і (в)). (е) $3 \cdot 2^{K-2} + 2^{n-1}$. Застосувати формулу включення-виключення (5.1) і результати попередніх задач. (є) 2^n. $|L \setminus T_1| = |L| - |L \cap T_1|$. (ж) 2^{n-1}. $|(T_0 \setminus T_1) \cap L| = |(T_0 \cap L) \setminus T_1| = |T_0 \cap L| - |T_0 \cap T_1 \cap L|$. (з) 2^{n-1}. $L \setminus (T_0 \cup T_1) = (L \setminus T_0) \setminus T_1$, далі див. (б), (в). (и) $2^{K-2} + 3 \cdot 2^{n-1}$. (і) $2^{K-1} - 2^n$, див. (г). (ї) $2^{K-1} + 2^n$. $P_2 \setminus (T_1 \setminus L) = (P_2 \setminus T_1) \cup L$, див. (г). **18.** Нехай $f(x_1, x_2, \dots, x_n) \in L^*$. Тоді існує така лінійна функція $g(x_1, x_2, \dots, x_n) = a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus a_{n+1}$, що $f = g^*$. Звідси $f(x_1, x_2, \dots, x_n) = \overline{a_1\overline{x}_1 \oplus a_2\overline{x}_2 \oplus \dots \oplus a_n\overline{x}_n \oplus a_{n+1}} =$

$= a_1(x_1 \oplus 1) \oplus a_2(x_2 \oplus 1) \oplus \dots \oplus a_n(x_n \oplus 1) \oplus a_{n+1} \oplus 1 = a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus (a_1 \oplus a_2 \oplus \dots \oplus a_n \oplus a_{n+1} \oplus 1)$, тобто $f \in L$. Навпаки, припустімо, що $f(x_1, x_2, \dots, x_n) \in L$, тобто $f(x_1, x_2, \dots, x_n) = a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus a_{n+1}$. Розглянемо лінійну функцію $g(x_1, x_2, \dots, x_n) = a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus (a_1 \oplus a_2 \oplus \dots \oplus a_n \oplus a_{n+1} \oplus 1)$. Неважко переконатися, що $g^* = f$, тому $f \in L^*$. **19.** (а), (б) — так. **20.** Із результату задачі 6.8.3 випливає, що для вектора значень $(a_0, a_1, \dots, a_k)$ самодвоїстої функції $f(x_1, x_2, \dots, x_n)$ виконується $a_i = \overline{a_{k-i}}$, $i = 0, 1, \dots, k$, $k = 2^n - 1$. Отже, $a_{m+j} = \overline{a_{m-j-1}}$, $j = 0, 1, \dots, m$. **21.** Твердження є наслідком результату попередньої задачі і тотожності $b + \overline{b} = 1$, $b \in B$. **22.** (а), (б) Для непарного n, бо $A^* = A \oplus (1 \oplus 1 \oplus \dots \oplus 1)$, і доданок у дужках містить $n+1$ одиницю. (в) Для непарного n, якщо $a = 0$, і для парного n, коли $a = 1$ (див. (а)). (г) За допомогою тотожності $a \vee b = ab \oplus a \oplus b$ дана формула перетворюється до рівносильної формули з пункту (б). **23.** (а) Спочатку зазначимо, що $\overline{f} \vee f^* \equiv 1$, бо в іншому разі $\overline{f} \equiv 1$ та $f^* \equiv 0$, і не виконується умова $|N_{\overline{f}}| = |N_{f^*}|$. На будь-якому наборі $(a_1, a_2, \dots, a_n)$ одна й тільки одна з функцій $\overline{f}$ або f^* набуває значення 1, тому $\overline{f}(a_1, a_2, \dots, a_n) = f^*(a_1, a_2, \dots, a_n)$ і $\overline{f} = f^*$. Якщо припустити, що існує такий набір $(a_1, a_2, \dots, a_n)$, для якого $\overline{f}(a_1, a_2, \dots, a_n) = 1$ і $f^*(a_1, a_2, \dots, a_n) = 1$, то на протилежному наборі $(\overline{a}_1, \overline{a}_2, \dots, \overline{a}_n)$ маємо $\overline{f}(\overline{a}_1, \overline{a}_2, \dots, \overline{a}_n) = 1$ і $f^*(\overline{a}_1, \overline{a}_2, \dots, \overline{a}_n) = 1$. А це означає, що $|N_{\overline{f}}| < |N_{f^*}|$, бо з $f(\overline{a_1, a_2, \dots, a_n}) = 0$ випливає $f^*(a_1, a_2, \dots, a_n) = 1$. (б) $f \oplus (f \wedge f^*) = f \wedge \overline{f^*} = \overline{\overline{f} \vee f^*}$, далі див. (а). **24.** 2^m, де $m = 2^{n-1}$. У векторі значень самодвоїстої функції можна довільно обирати значення (0 чи 1) лише для половини його координат (див. 6.11.20). **25.** Позначимо $K = 2^n$ і $m = 2^{n-1}$. (а) 2^{m-1}. (б) $2^{K-1} + 2^{m-1}$. (в), (г), (д) 2^{m-1}. (е) 2^n. Лінійна функція $a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus a_{n+1}$ самодвоїста тоді й тільки тоді, коли кількість її коефіцієнтів $a_1, a_2, \dots, a_n$, що дорівнюють 1, непарна (див. 6.11.18 та 6.11.17(в)). (є) $2^{m-1} + 2^{n-1}$. (ж) $2^{K-1} + 2^{m-1} + 2^{n-1}$. Застосувати формулу включення-виключення (5.1) і результати попередніх задач. (з) $2^{m-1} + 2^{n-1}$. (и), (і) 0. (ї) 2^n. **26.** (а) Слід довести, що булева функція $f(x_1, x_2, \dots, x_n)$ належить кожній із даних множин тоді й тільки тоді, коли вона лінійна і в її поліномі Жегалкіна $a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus a_{n+1}$ вільний член $a_{n+1} = 0$, а число одиничних коефіцієнтів серед $a_1, a_2, \dots, a_n$ непарне. (б) Функція належить будь-якій із цих множин тоді й тільки тоді, коли вона самодвоїста і зберігає константу 0 (що рівносильно тому, що вона самодвоїста і зберігає константу 1). **27.** Довести, що для довільних множин булевих функцій A і B виконуються такі рівності: $(A \cup B)^* = A^* \cup B^*$, $(A \cap B)^* = A^* \cap B^*$ і $(A \setminus B)^* = A^* \setminus B^*$. Відтак, використати, що множини C й L самодвоїсті (див. 6.11.18), а $T_0^* = T_1$ і $T_1^* = T_0$ (див. 6.11.9(а)). **28.** Якщо $f \notin M$, то існують такі набори $\boldsymbol{c}$ і $\boldsymbol{d}$, що $\boldsymbol{c} < \boldsymbol{d}$ і $f(\boldsymbol{c}) > f(\boldsymbol{d})$. У разі $\rho(\boldsymbol{c}, \boldsymbol{d}) = 1$ твердження справедливе. Припустімо, що $\rho(\boldsymbol{c}, \boldsymbol{d}) = k$ і $k > 1$. Розглянемо послідовність сусідніх наборів $\boldsymbol{h}_0, \boldsymbol{h}_1, \ldots, \boldsymbol{h}_k$, де $\boldsymbol{h}_0 = \boldsymbol{c}$, $\boldsymbol{h}_k = \boldsymbol{d}$ і $\boldsymbol{h}_{i-1} < \boldsymbol{h}_i$, $i = 1, 2, \ldots, k$. Оскільки $f(\boldsymbol{h}_0) = 1$ і $f(\boldsymbol{h}_k) = 0$, то існує j, для якого $f(\boldsymbol{h}_{j-1}) > f(\boldsymbol{h}_j)$, $j = 1, 2, \ldots, k$. Тоді $\boldsymbol{a} = \boldsymbol{h}_{j-1}$ і $\boldsymbol{b} = \boldsymbol{h}_j$ — шукані набори. **29.** Перевірити обидва співвідношення для наборів $(0, a_2, \ldots, a_n)$ і $(1, a_2, \ldots, a_n)$ та скористатися тим, що нерівність $c \le d$ рівносильна $c \vee d = d$ і $c \wedge d = c$ для $c, d \in B$. **30.** (а), (в), (г) — монотонні. **31.** Застосувати метод математичної індукції за числом змінних, за якими здійснюється розклад за допомогою формул із задачі 6.11.29. **32.** (а) 0. (б) 2. (в) $n + 2$. $M \cap L = \{x_1, x_2, \ldots, x_n, 0, 1\}$. (г) n. (д) $2^n - 2$. Даній множині належать лінійні функції, у поліномі Жегалкіна $a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n \oplus a_{n+1}$ яких число одиничних коефіцієнтів серед $a_1, a_2, \ldots, a_n$ парне і не дорівнює нулю (див. 6.11.25(е)). **33.** Якщо $f(x_1, x_2, \ldots, x_n) \in M$, то для будь-яких $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ і $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ таких, що $\boldsymbol{a} \le \boldsymbol{b}$, виконується нерівність $f(\boldsymbol{a}) \le f(\boldsymbol{b})$. Позначимо $\bar{\boldsymbol{a}} = (\bar a_1, \bar a_2, \ldots, \bar a_n)$ і $\bar{\boldsymbol{b}} = (\bar b_1, \bar b_2, \ldots, \bar b_n)$. Тоді $\boldsymbol{a} \le \boldsymbol{b}$ рівносильно $\bar{\boldsymbol{b}} \le \bar{\boldsymbol{a}}$, $f^*(\bar{\boldsymbol{b}}) = \bar f(\boldsymbol{b})$, $f^*(\bar{\boldsymbol{a}}) = \bar f(\boldsymbol{a})$, тому $f^*(\bar{\boldsymbol{b}}) \le f^*(\bar{\boldsymbol{a}})$. Отже, $f^* \in M$. Усі наведені міркування оборотні, тому $f \in M$ тоді й тільки тоді, коли $f^* \in M$. **34.** Нехай f — монотонна функція. Якщо $f \notin T_0$, то $f \equiv 1$, тобто $f \in T_1$ і $f \in L$. Якщо $f \notin T_1$, то $f \equiv 0$ і $f \in T_0$, $f \in L$. Якщо $f \notin L$, то $f \ne \text{const}$, отже $f \in T_0$, $f \in T_1$. **35.** Розглянемо поліном Жегалкіна $a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n \oplus a_{n+1}$ довільної лінійної функції $f(x_1, x_2, \ldots, x_n)$. Якщо $a_{n+1} = 0$, то $f \in T_0$. Якщо $a_{n+1} = 1$ і число k коефіцієнтів серед $a_1, a_2, \ldots, a_n$, які дорівнюють 1, парне, то $f \in T_1$. Якщо ж число k непарне, то $f \in C$ (див. 6.11.18, 6.11.25(е), а також 6.11.32(в)).

6.12. 1. (а) Ні. $\Sigma \subseteq L$. (б) Ні. $\Sigma \subseteq T_0$. (в) Ні. $\Sigma \subseteq T_1$. (г) Так. **2.** (а) Ні. $\Sigma \subseteq T_0$. (б), (в) Так. **3.** (а), (в) Ні. **4.** (а) $\{0, x \to y\}$, $\{x \oplus y, x \to y\}$. (б) $\{x \oplus y, x \to y\}$. **5.** (а) Повна. (б) Неповна. $g \equiv 0$ й $f \equiv 0$. **6.** (а) Повна. Наприклад, $\{m(x, y, z), \bar x, x \oplus y \oplus 1\} \subseteq \Sigma$. (б) Неповна. $\Sigma \subseteq T_1$. (в) Неповна. $\Sigma \subseteq L$. **7.** З $f \notin T_0 \cup T_1$ випливає $f \notin M$. Якщо припустити, що $f \in L$, то з урахуванням $f \notin T_0 \cup T_1$, дістали б, що $f \in C$ (див. 6.11.25(е)), а це суперечить умові. **8.** (а) $\{\,|\,\}$, $\{\downarrow\}$. (б) $\{\oplus, \to\}$, $\{\sim, \not\leftarrow\}$, $\{\to, \not\to\}$.

(в) $\{\oplus, \sim, \wedge\}$, $\{\oplus, \sim, \vee\}$. **9.** Ні. Побудувати критеріальну таблицю для всіх таких функцій і переконатися, що не існує інших базисів, окрім наведених у попередній задачі.

6.13. 2. (а) (1011). (б) (00101111). **4.** Перетворити спочатку A до рівносильної формули, яка має k операндів. (а) $(a + b) * c$. (б) $\neg a + \neg b + \neg c$. (в) $(\neg a + b) * (a + c)$. (г) $(a + b) * (\neg a + c)$. **5.** Відповідна формула — $A = k_1 * (k_2 + k_3)$, де k_1 — контакт голови комітету. **6.** (а) $k_1 + k_2 + + k_3 + k_4$. (б) $k_1 * (k_2 + k_3 + k_4) + k_2 * (k_3 + k_4) + k_3 * k_4$. (в) $k_1 * k_2 * k_3 * k_4$. (г) $k_1 * k_2 * (k_3 + k_4) + k_3 * k_4 * (k_1 + k_2)$. **7.** $k_1 * k_2 * k_3 + \neg k_2 * \neg k_4$.

6.14. 1. (а) $(\bar x \to \bar y) \to \overline{x \to y}$. (б) $\overline{\bar{\bar x} \wedge y \wedge \bar{\bar y} \wedge z}$ (в) Визначити відповідний поліном Жегалкіна. (г) $((x \vee y \vee z \vee (w \oplus 1)) \oplus 1) \vee ((x \vee y \vee (z \oplus 1) \vee w) \oplus 1) \vee (((x \oplus 1) \vee y \vee (z \oplus 1) \vee (w \oplus 1)) \oplus 1) \vee (((x \oplus 1) \vee (y \oplus 1) \vee (z \oplus 1) \vee w) \oplus 1)) \oplus 1$, бо $a \wedge b = ((a \oplus 1) \vee (b \oplus 1)) \oplus 1$. **2.** Визначити ДДНФ і максимально спростити її за допомогою рівносильних перетворень. **3.** Перетворити A до простішої рівносильної формули. (а) $a(\bar b \vee d) \vee bcd$. (б) $a \vee b(f \vee \bar c\bar e\bar f)$. **4.** Справедливість твердження випливає з ізоморфізму алгебри РК схем R і алгебри булевих формул D'_B. **5.** Нехай функція $f(x_1, x_2, \ldots, x_n)$ набуває значення 1 не більше ніж на половині всіх наборів. Тоді її ДДНФ містить не більше $2^{n-1} - 1$ диз'юнкцій і не більше $(n-1)2^{n-1}$ кон'юнкцій, тобто загальна кількість бінарних операцій у ДДНФ не перевищує $n2^{n-1} - 1$. Якщо ж $|N_f| > 2^{n-1}$, то ДКНФ функції f міститиме не більше ніж $n(2^{n-1} - 1) - 1$ диз'юнкцій і кон'юнкцій.

6.15. 1. (а) $\{x, xy, y\bar z\}$. (б) $\{x\bar y\}$. (в) $\{x\bar z, y\bar z, xyzw\}$. **2.** (а) $\{x\bar y, y\bar z\}$. (б) $\{y\bar z, x\bar z, xy, \bar x\bar y z\}$. **3.** Припустімо, що проста імпліканта W монотонної функції $f(x_1, x_2, \ldots, x_n)$ містить заперечення якоїсь змінної x_i. Тоді $W = \bar x_i V$, де V — елементарна кон'юнкція. Розглянемо множину N_W двійкових наборів довжини n, на яких W (і функція f) набуває значення 1. Нехай $(a_1, a_2, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) \in N_W$. Тоді внаслідок монотонності функції f матимемо $f(a_1, a_2, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n) = 1$. Це означатиме, що x_iV — також імпліканта функції f. Якщо $\bar x_iV$ і x_iV — імпліканти f, то імплікантою f буде й V, що суперечитиме припущенню про простоту W. **4.** 2^S, де $S = 2^n - 2^{n-k}$, бо $f(x_1, x_2, \ldots, x_n)$ має набувати значення 1 на всіх 2^{n-k} наборах, у яких k координат, що відповідають змінним імпліканти K, мають певні фіксовані значення. **5.** Застосувати метод математичної індукції за числом змінних n і формулу (6.6) розкладу булевої функції за однією змінною. **6.** Нехай $K_1 = x_1^{a_1}x_2^{a_2}\ldots x_n^{a_n}$ і $K_2 = x_1^{b_1}x_2^{b_2}\ldots x_n^{b_n}$. Позначимо через $S(K_1, K_2)$ елементарну кон'юнкцію, утворену з K_1 викресленням тих букв $x_i^{a_i}$, для яких $a_i \neq b_i$. Можна довести, що довільну імпліканту (зокрема, й просту) функції f можна зобразити у вигляді $S(K, L)$, де K і L — певні елементарні кон'юнкції з ДДНФ функції f. **7.** Для доведення необхідності застосувати індукцію за числом змінних імпліканти K. **8.** Див. попередні задачі.

6.16. 1. (а) $x \vee \bar{y}$. (б) $\bar{x}\bar{z} \vee \bar{y}\bar{z} \vee xyz$. (в) $\bar{x}\bar{y}z \vee xz$. **2.** (а) $x \vee y \vee \bar{z}$. (б) $\bar{x}\bar{y}\bar{z} \vee yz$. **3.** (а) $\bar{x}\bar{y} \vee xy \vee xz$. (б) $\bar{x}\bar{z} \vee \bar{y}\bar{z} \vee xy$. (в) $\bar{x}\bar{y}\bar{z}w \vee \bar{x}yw \vee \vee \bar{y}zw \vee \bar{x}yz \vee y\bar{z}\bar{w} \vee x\bar{y}zw \vee x\bar{y}z \vee xy\bar{w}$.

6.18. 1. (а) $\bar{y}z \vee y\bar{z} \vee xy$, $\bar{y}z \vee y\bar{z} \vee xz$ — тупикові та мінімальні ДНФ. (б) $\bar{x}z \vee x\bar{y}$. (в) $\bar{x}z\bar{w} \vee xw \vee \bar{z}w \vee \bar{x}\bar{y}\bar{z}$, $\bar{x}z\bar{w} \vee xw \vee \bar{z}w \vee \bar{x}\bar{y}\bar{w}$ — тупикові і мінімальні ДНФ. **2.** (а) $\bar{x}z$, $x\bar{y}$, $\bar{y}zw$. (б) $xy\bar{z}$, $\bar{x}yz$. (в) xw, $\bar{y}z$, $y\bar{w}$, $\bar{x}yzw$. **3.** (б), (в) Тупикова, мінімальна і найкоротша ДНФ. **4.** 58 тупикових і 12 мінімальних.

6.19. 1. (а) $\bar{x}y \vee \bar{y}z$. (б) $\bar{x}\bar{y} \vee \bar{x}\bar{z} \vee \bar{y}z$. (в) $\bar{x}z \vee zw \vee yw \vee \bar{y}z \vee xy$. **2.** (а) $\bar{x}\bar{y}\bar{z} \vee yz \vee xz$. (б) $xy\bar{w} \vee \bar{x}\bar{z}w \vee \bar{x}yz \vee \bar{y}zw \vee x\bar{y}z \vee y\bar{z}w \vee \bar{x}zw \vee \bar{x}yw$. **3.** (а) $\bar{x}z \vee xy$. (б) $\bar{y}z \vee x\bar{z}$ (в) $\bar{y}zw \vee \bar{x}zw \vee yw \vee xy\bar{z} \vee x\bar{z} \vee x\bar{y}w$.

7. Теорія автоматів

7.1. 1. Покладемо x_1 — логічний зсув праворуч, x_2 — арифметичний зсув праворуч, x_3 — циклічний зсув праворуч. Станами автомата Z будуть усі вісім можливих варіантів трирозрядного двійкового коду. Суміщена таблиця автомата Z має такий вигляд.

δ/λ	000	001	010	011	100	101	110	111
x_1	000/0	000/1	001/0	001/1	010/0	010/1	011/0	011/1
x_2	000/0	000/1	001/0	001/1	110/0	110/1	111/0	111/1
x_3	000/0	010/0	100/0	110/0	001/1	011/1	101/1	111/1

2. Суміщена таблиця автомата P має такий вигляд.

δ/λ	000	001	010	011	100	101	110	111
0	000/0	010/0	100/0	110/0	000/1	010/1	100/1	110/1
1	001/0	011/0	101/0	111/0	001/1	011/1	101/1	111/1

3. Позначимо початковий стан автомата A через П. Вихідний сигнал 1 відповідатиме переходу в стан T, а 0 — переходу в стан H.

δ/λ	П	Т	Н
x_1	Т/1	Т/1	Н/0
x_2	Н/0	Т/1	Н/0
x_3	Н/0	Н/0	Н/0

4. Позначимо через П початковий стан автомата B, через Ц — стан, у якому автомат B знаходиться, якщо на його вхід подано послідовність цифр, що відповідає цілій частині дійсного числа (зі знаком або без знаку), через Д — стан, що відповідає дробовій частині числа, через Н — стан, що відповідає послідовності символів, яка не є числовою константою. Вихідний сигнал автомата B дорівнює 0 тоді й лише тоді, коли автомат переходить у стан Н.

δ/λ	П	Ц	Д	Н
x_1	Ц/1	Н/0	Н/0	Н/0
x_2	Ц/1	Ц/1	Д/1	Н/0
x_3	Д/1	Д/1	Н/0	Н/0
x_4	Н/0	Н/0	Н/0	Н/0

5. Позначимо через 0 стан автомата затримки, який "запам'ятовує", що в попередній момент часу на його вхід було подано сигнал 0, а через 1 — стан, що "запам'ятовує" вхідний сигнал 1.

δ/λ	0	1
0	0/0	0/1
1	1/0	1/1

6. Позначимо через 0 сигнал, що відповідає очікуванню бажаної четвірки символів, 1 — стан, що фіксує появу на вході автомата першого 0 четвірки, 2 — стан, що відповідає появі пари 01, 3 — стан, що відповідає появі 011, і 4 — стан, у який автомат переходить, коли останні вхідні сигнали — 0110.

δ/λ	0	1	2	3	4
0	1/0	1/0	1/0	4/1	1/0
1	0/0	2/0	3/0	0/0	2/0

7. Розглянемо суміщену таблицю "генератора парності". Він потрапляє у стан 3, якщо кількість одиниць серед перших двох символів чергової трійки непарна, у стан 4 — якщо ця кількість парна. Автомат опиняється у стані 5, якщо вхідна трійка символів містить непарну кількість одиниць, і в стані 6 — якщо ця кількість парна.

δ/λ	0	1	2	3	4	5	6
0	1/0	4/0	3/0	5/0	6/0	0/0	0/1
1	2/1	3/1	4/1	6/1	5/1	0/0	0/1

8. Покладемо $x_1 = (0, 0)$, $x_2 = (0, 1)$, $x_3 = (1, 0)$, $x_4 = (1, 1)$. Позначимо через 0 початковий стан і одночасно стан, що відповідає рівності обох даних чисел, а через 1 — стан, що відповідає ситуації "перше з чисел є більшим", через 2 — стан, що відповідає ситуації, коли друге число більше першого.

δ/λ	0	1	2
x_1	0/0	1/0	2/0
x_2	2/1	1/0	2/1
x_3	1/1	1/1	2/0
x_4	0/1	1/1	2/1

9. Позначимо через П початковий стан автомата, через Т — стан, що фіксує правильну відповідність дужок у вхідному слові, а через Н — стан, що свідчить про помилку в запису дужок у виразі. Стан k відповідає ситуації наявності k незакритих лівих дужок, $k = 1, 2, \ldots, n$.

δ/λ	П	1	2	3	…	$n-1$	n	Т	Н
x_1	1/0	2/0	3/0	4/0	…	n/0	Н/0	Т/1	Н/0
x_2	Н/0	П/0	1/0	2/0	…	$n-2$/0	$n-1$/0	Т/1	Н/0
x_3	П/0	1/0	2/0	3/0	…	$n-1$/0	n/0	Т/1	Н/0
x_4	Т/1	Н/0	Н/0	Н/0	…	Н/0	Н/0	Т/1	Н/0

10. Позначимо через Г вихідний сигнал, що відповідає руху вгору, через Н — руху вниз і через С — ситуації, коли ліфт не рухається.

δ/λ	1	2	3	4
1	1/С	1/Н	1/Н	1/Н
2	2/Г	2/С	2/Н	2/Н
3	3/Г	3/Г	3/С	3/Н
4	4/Г	4/Г	4/Г	4/С

7.2. 1. 1) Для $x \in X$ $\lambda(a_i, x)$ визначають за таблицею виходів автомата. 2) Для довільного слова $p \in X^*$ і довільного вхідного сигналу $x \in X$ $\lambda(a_i, px) = \lambda(\delta(a_i, p), x)$. **2.** 1) Першу властивість доведемо методом математичної індукції за довжиною вхідного слова p. База індукції: для $|p| = 0$ маємо $p = e$ та $|\varphi_A(p)| = |e| = 0$. Якщо $|p| = 1$, тобто $p = x_j$ і $x_j \in X$, то $\varphi_A(p) = \varphi_A(x_j) = \varphi_A(ex_j) = \varphi_A(e)\lambda(a_1, x_j) = e\lambda(a_1, x_j) = \lambda(a_1, x_j) = y_k$ і $y_k \in Y$, тому $|p| = |\varphi_A(p)|$. Індукційний крок: нехай дана рівність справджується для всіх слів $p \in X^*$ довжини n, $n \geq 1$. Розглянемо довільне слово $p' = px_j$ довжини $n + 1$, де $|p| = n$. Тоді $\varphi_A(p') = \varphi_A(px_j) = \varphi_A(p)\lambda(a_1, px_j)$. Звідси $|\varphi_A(p')| = |\varphi_A(px_j)| = |\varphi_A(p)| + |\lambda(a_1, px_j)| = |p| + 1 = |p'|$. Першу властивість доведено. 2) Для доведення другої властивості розглянемо слово $p_2 = x_{i1}x_{i2}\ldots x_{ik}$. Тоді $\varphi_A(p_1p_2) = \varphi_A(p_1x_{i1}x_{i2}\ldots x_{ik}) = \varphi_A(p_1x_{i1}x_{i2}\ldots x_{ik-1})$ $\lambda(a_1, p_1x_{i1}x_{i2}\ldots x_{ik}) = \varphi_A(p_1x_{i1}x_{i2}\ldots x_{ik-2})\lambda(a_1, p_1x_{i1}x_{i2}\ldots x_{ik-1})\ \lambda(a_1, p_1x_{i1}x_{i2}\ldots x_{ik}) = \ldots = \varphi_A(p_1)\lambda(a_1, p_1x_{i1})\lambda(a_1, p_1x_{i1}x_{i2})\ldots\lambda(a_1, p_1x_{i1}x_{i2}\ldots x_{ik}) = q_1q_2$. За умовою $|p_1| = |q_1|$, а з першої властивості маємо $|p_1| = |\varphi_A(p_1)|$. Отже, $|\varphi_A(p_1)| = |q_1|$, тому з урахуванням отриманого вище співвідношення доходимо висновку, що $\varphi_A(p_1) = q_1$. **3.** Доведення проведемо методом математичної індукції за довжиною слова p_2. База індукції. Для $|p_2| = 0$, тобто $p_2 = e$ маємо $\varphi_A^i(p_1p_2) = \varphi_A^i(p_1e) = \varphi_A^i(p_1) = \varphi_A^i(p_1)e = \varphi_A^i(p_1)\varphi_A^{i1}(e) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)$. Якщо $|p_2| = 1$, тобто $p_2 = x_j$ і $x_j \in X$, то $\varphi_A^i(p_1p_2) = \varphi_A^i(p_1x_j) = \varphi_A^i(p_1)\lambda(a_1, p_1x_j) = \varphi_A^i(p_1)\lambda(\delta(a_i, p_1), x_j) = \varphi_A^i(p_1)\lambda(a_{i1}, x_j) = \varphi_A^i(p_1)\varphi_A^{i1}(x_j) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)$. Окрім того, виконується співвідношення $\delta(a_1, p_1p_2) = \delta(a_1, p_1x_j) = \delta(\delta(a_1, p_1), x_j) = \delta(\delta(a_1, p_1), p_2)$. Індукційний крок. Припустімо, що для всіх слів p_2 довжини n ($n \geq 1$) виконується рівність $\varphi_A^i(p_1p_2) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)$, де $a_{i1} = \delta(a_1, p_1)$, а також, що має місце співвідношення $\delta(a_1, p_1p_2) = \delta(\delta(a_1, p_1), p_2)$. Розглянемо довільне вхідне слово $p = p_1p_2'$ для $|p_2'| = n + 1$, тобто $p_2' = p_2x$, $p_2 \in X^*$, $|p_2| = n$ та $x \in X$. Тоді послідовно матимемо $\varphi_A^i(p_1p_2') = \varphi_A^i(p_1p_2x) = \varphi_A^i(p_1p_2)\lambda(a_i, p_1p_2x) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)\lambda(\delta(a_i, p_1p_2), x) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)\lambda(\delta(\delta(a_1, p_1), p_2), x) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)\ \lambda(\delta(a_{i1}, p_2), x) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2)\lambda(a_{i1}, p_2x) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2x) = \varphi_A^i(p_1)\varphi_A^{i1}(p_2')$. **4.** (а) 01011. (б) 100000. (в) 010100011. **5.** (а) 00000. (б) 101001. (в) 010100010. **6.** Доведення проведемо методом математичної індукції за довжиною слова w. 1) Для $|w| = 1$, тобто $w = x_{i_1}$ та $x_{i_1} \in X$ маємо $M(w) = M(x_{i_1})$, бо $\delta(a_i, w) = \delta(a_i, x_{i_1})$. 2) Припустимо, що твердження виконується для всіх слів w довжини n ($n \geq 1$). Розглянемо довільне вхідне слово $w' = wx_l$ довжини $n + 1$, $w = x_{i_1}x_{i_2}\ldots x_{i_n}$ і $x_l \in X$. Доведемо, що $M(w') = M(w)M(x_l)$. Елемент $m_{ij}(w')$ матриці $M(w')$ дорівнює 1 тоді й тільки тоді, коли $a_j = \delta(a_i, w')$. Відповідний елемент матриці $M(w)M(x_l)$ можна обчислити за формулою $p_{ij} = m_{i1}(w)m_{1j}(x_l) + m_{i2}(w)m_{2j}(x_l) + \ldots + m_{is}(w)m_{sj}(x_l)$, де $s = |U|$. Із властивостей матриці переходів випливає, що остання сума дорівнює 1 тоді й лише тоді, коли існує значення k, для якого $m_{ik}(w)m_{kj}(x_l) = 1$, що рівносильно $m_{ik}(w) = 1$ й $m_{kj}(x_l) = 1$ або $a_k = \delta(a_i, w)$ і $a_j = \delta(a_k, x_l)$. Із двох останніх співвідношень маємо $a_j = \delta(\delta(a_i, w), x_l) = \delta(a_i, wx_l) = \delta(a_i, w')$. Отже, $p_{ij} = 1$ тоді й лише тоді, коли $a_j = \delta(a_i, w')$, тобто $m_{ij}(w') = p_{ij}$. Враховуючи припущення індукції, отримаємо $M(w') = M(w)M(x_l) = M(x_{i_1})M(x_{i_2}) \cdots M(x_{i_n})M(x_l)$. **7.** Твердження можна довести методом математичної індукції за довжиною вхідного слова w (див. попередню задачу).

7.3. 1. Відношення невідрізнюваності станів означено через рівність відповідних відображень, а відношення рівності рефлексивне, симетричне і транзитивне, тобто є еквівалентністю. Рефлексивність відношення невідрізнюваності автоматів випливає з того, що будь-який стан автомата невідрізнюваний від самого себе. Властивості симетричності і транзитивності є безпосередніми наслідками означення відношення невідрізнюваності автоматів, а також симетричності і транзитивності відношення рівності. **2.** Нехай γ — ізоморфізм автоматів A_1 і A_2. Для доведення твердження слід довести, що для будь-якого стану a автомата A_1 стани a та $\gamma(a)$ невідрізнювані (див. теорему 7.1). Справедливість даного твердження випливає також із того, що ізоморфізм є окремим випадком гомоморфізму. **3.** Наприклад, автомат $A = (X, Y, U_1, \delta_1, \lambda_1)$, де $X = \{x\}$, $Y = \{y\}$, $U_1 = \{1\}$, $\delta_1(1, x) = 1$, $\lambda_1(1, x) = y$ і автомат $B = (X, Y, U_2, \delta_2, \lambda_2)$, де $U_2 = \{1, 2\}$, $\delta_2(1, x) = 2$, $\delta_2(2, x) = 1$, $\lambda_2(1, x) = \lambda_2(2, x) = y$. **4.** Нехай γ_1 і γ_2 — гомоморфізми автомата A_1 на A_2 і на A_3 відповідно. Для доведення твердження достатньо показати, що для довільного стану a автомата A_2 стани a і $b = \gamma_2(c)$, $c \in \gamma_1^{-1}(a)$ невідрізнювані (див. теорему 7.1). З іншого боку, із теореми 7.1 випливає, що автомати A_1 і A_2, а також A_1 і A_3 невідрізнювані. Із властивостей відношення невідрізнюваності автоматів (див. задачу 7.3.1) отримаємо невідрізнюваність автоматів A_2 і A_3.

7.4. 1. Нехай $Z = (X, Y, Q, \delta', \lambda')$ і $Z' = (X, Y, Q', \delta'', \lambda'')$. Друга властивість $\lambda'(b, x) = \lambda''(\eta(b), x)$ випливає з того, що $\lambda'(b, x) = \varphi_Z^b(x)$, $\lambda''(\eta(b), x) = \varphi_{Z'}^{b'}(x)$, $b' = \eta(b)$, і з того, що стани b і b' невідрізнювані. Стани b і $\eta(b)$ невідрізнювані, тому для всіх $x \in X$ і $p \in X^*$ маємо $\varphi_Z^b(xp) = \varphi_{Z'}^{b'}(xp)$, $b' = \eta(b)$. Звідси $\lambda'(b, x)\varphi_Z^{\delta'(b, x)}(p) = \lambda''(b', x)\varphi_{Z'}^{\delta''(b', x)}(p)$, тобто $\varphi_Z^{\delta'(b, x)}(p) = \varphi_{Z'}^{\delta''(b', x)}(p)$. Це означає, що стан $\delta'(b, x)$ автомата Z і стан $\delta''(b', x)$ автомата Z' невідрізнювані. Отже, виходячи з означення відповідності η, матимемо $\eta(\delta'(b, x)) = \delta''(\eta(b), x)$. **2.** Розглянемо мінімальні та невідрізнювані автомати $Z = (X, Y, Q, \delta', \lambda')$ і $Z' = (X, Y, Q', \delta'', \lambda'')$, які належать одному класу K. Із доведення теореми 7.2 випливає, що існує бієктивна відповідність η між множинами станів Q і Q' така, що пари $(b, b') \in \eta$ складаються з невідрізнюваних станів. Згідно з попередньою задачею для η виконуються умови (7.6), тому η є ізоморфізмом для автоматів Z і Z'. **3.** Використовуючи метод доведення від супротивного, припустімо, що автомат $A = (X, Y, U, \delta, \lambda)$ не є мінімальним. Тоді серед його станів є невідрізнювані. Застосуємо до автомата A процедуру, описану в доведенні теореми 7.2, і побудуємо відповідний автомат $Z = (X, Y, Q, \delta', \lambda')$. Оскільки $Z \in K$ і $|Q| = |U/N| < |U|$, то отримуємо суперечність стосовно припущення, що автомат A має найменшу можливу кількість станів серед усіх автоматів класу K.

7.5. 1. Для доведення леми 7.1 розглянемо довільне вхідне слово p довжини l, $l < k$. Додавши до нього будь-яке слово r довжини $k - l$, за умовою леми отримаємо $\varphi_A^i(pr) = \varphi_A^j(pr)$. Звідси за формулою (7.4) $\varphi_A^i(pr) = \varphi_A^i(p)\varphi_A^{i_1}(r)$ і $\varphi_A^j(pr) = \varphi_A^j(p)\varphi_A^{j_1}(r)$, де $a_{i_1} = \delta(a_i, p)$ й $a_{j_1} = \delta(a_j, p)$. Виходячи з умов автоматності для відображення φ_A, отримаємо $|\varphi_A^i(p)| = |p| = |\varphi_A^j(p)|$, отже $\varphi_A^i(p) = \varphi_A^j(p)$. Справедливість леми 7.2 випливає з того, що кожне вхідне слово $p \in X^*$ має певну довжину k, $k = 0, 1, 2, \ldots$. Властивості еквівалентності для k-невідрізнюваності можна обгрунтувати подібно до того, як це було зроблено для відношення невідрізнюваності станів (див. задачу 7.3.1). **2.** Доведення проведемо методом математичної індукції. Нехай $k = 1$ та $Q_1 \neq Q_2$. Припустимо, що $|Q_1| = 1$. Це означає, що будь-які стани a_i й a_j автомата A, а також стани $a_{i_1} = \delta(a_i, x)$ і $a_{j_1} = \delta(a_j, x)$ для всіх $x \in X$ є 1-невідрізнюваними. За лемою 7.4 стани a_i і a_j будуть 2-невідрізнювані, тому $Q_1 = Q_2$, що суперечить умові. Отже, $|Q_1| \geq 2 = k + 1$. Базу індукції доведено. Припустимо, що твердження справедливе для $k = n$ $(n \geq 1)$ і виконується $|Q_{n+1}| \neq |Q_{n+2}|$. З останньої нерівності випливає, що $|Q_n| \neq |Q_{n+1}|$, бо за теоремою 7.3 рівність $|Q_n| = |Q_{n+1}|$ тягне $|Q_{n+1}| = |Q_{n+2}|$. Отже, за припущенням індукції маємо $|Q_{n+1}| \geq n + 1$. Як було показано вище, класи Q_{n+1} є підкласами Q_n, тому якщо $Q_n \neq Q_{n+1}$, то $|Q_{n+1}| > |Q_n|$. Отже, $|Q_{n+1}| > |Q_n| \geq n + 1$, тобто $|Q_{n+1}| \geq n + 2$. **3.** Якщо таке слово p існує, то відповідні стани, очевидно, відрізнювані. Для доведення оберненого твердження припустимо, що для деяких відрізнюваних станів a_i та a_j автомата $A = (X, Y, U, \delta, \lambda)$ не існує вхідного слова p довжини не більше $|U| - 1$, для якого $\varphi_A^i(p) \neq \varphi_A^j(p)$. Це означатиме, що стани a_i й a_j належать одному класу k-невідрізнюваності для $k = 1, 2, \ldots, |U| - 1$. Водночас з існування відрізнюваних станів a_i та a_j випливає, що $Q_k \neq Q_{k+1}$ для всіх $k = 1, 2, \ldots, |U| - 1$ (в іншому разі, за теоремою 7.3 мали б бути невідрізнюваними всі стани, що опинилися в одному класі). Тоді з твердження попередньої задачі випливає нерівність $|Q_{s-1}| \geq s$ для $s = |U| - 1$, яка свідчить, що розбиття Q_{s-1} складається з одноелементних підмножин множини U, тому стани a_i й a_j не можуть належати тому самому класу. Отримана суперечність і доводить твердження. **4.** Припустимо, що $Q_{s-1} \neq Q_s$, тоді $|Q_{s-1}| \geq s$ (див. задачу 7.5.2). Отже, класи розбиття Q_{s-1} є одноелементними підмножинами множини станів U, і Q_s уже не може бути підрозбиттям для Q_{s-1}, тобто $Q_{s-1} = Q_s$. **5.** Якщо стани a_i та a_j невідрізнювані, то вони, очевидно, будуть k-невідрізнювані й для $k = |U| - 1$. Навпаки, нехай стани a_i й a_j автомата $A = (X, Y, U, \delta, \lambda)$ k-невідрізнювані для $k = |U| - 1$. Із твердження попередньої задачі і теореми 7.3 випливає, що ($|U| - 1$)-невідрізнювані стани будуть взагалі невідрізнюваними. **6.** Суміщена таблиця шуканого автомата має такий вигляд:

δ/λ	a_1	a_2	a_3	a_4	…	a_{n-1}	a_n
0	$a_2/0$	$a_3/0$	$a_4/0$	$a_5/0$	…	$a_n/0$	$a_n/0$
1	$a_1/1$	$a_1/0$	$a_2/0$	$a_3/0$	…	$a_{n-2}/0$	$a_{n-1}/0$

Стани a_{n-1} та a_n є $(n-2)$-невідрізнюваними, бо для будь-якого вхідного слова p довжини $n - 2$ $\varphi_A^{n-1}(p) = \varphi_A^n(p) = 00\ldots0$. Водночас для слова $w = 11\ldots1$ довжини $n - 1$ $\varphi_A^{n-1}(w) = 00\ldots01$ і $\varphi_A^n(w) = 00\ldots00$. **7.** Припустимо, що стани a_i та a_j автомата $A = (X, Y, U, \delta, \lambda)$ невідрізнювані. Розглянемо довільні вхідні слова w, $w_1 \in X^*$. Тоді $\varphi_A^i(ww_1) = \varphi_A^j(ww_1)$. За формулою (7.4) $\varphi_A^i(w)\varphi_A^{i_1}(w_1) = \varphi_A^j(w)\varphi_A^{j_1}(w_1)$, де $a_{i_1} = \delta(a_i, w)$ і $a_{j_1} = \delta(a_j, w)$. Звідси з урахуванням рівності $\varphi_A^i(w) = \varphi_A^j(w)$ матимемо $\varphi_A^{i_1}(w_1) = \varphi_A^{j_1}(w_1)$. Отже, стани a_{i_1} і a_{j_1} невідрізнювані. Навпаки, нехай стани $a_{i_1} = \delta(a_i, w)$ і $a_{j_1} = \delta(a_j, w)$ невідрізнювані для довільного $w \in X^*$.

Зокрема, ця умова виконується і для $w = e$, тобто стани $a_i = \delta(a_i, e)$ і $a_j = \delta(a_j, e)$ невідрізнювані. **8.** Обґрунтуванням даного алгоритму є результат попередньої задачі: якщо для станів a_i і a_j автомата $A = (X, Y, U, \delta, \lambda)$ існує вхідне слово $w \in X^*$, для якого стани $a_{i_1} = \delta(a_i, w)$ й $a_{j_1} = \delta(a_j, w)$ відрізнювані, то стани a_i і a_j також відрізнювані і належать різним класам невідрізнюваності. **9.** (а) $b_1 = \{a_1, a_6\}$, $b_2 = \{a_2, a_4, a_5, a_7\}$, $b_3 = \{a_3\}$. (б) $b_1 = \{a_1, a_3, a_4\}$, $b_2 = \{a_2, a_5\}$, $b_3 = \{a_6\}$, $b_4 = \{a_7\}$. Відповідні суміщені таблиці мінімальних автоматів мають такий вигляд:

(а)

δ/λ	b_1	b_2	b_3
x_1	$b_2/1$	$b_3/1$	$b_1/1$
x_2	$b_1/2$	$b_2/2$	$b_3/2$
x_3	$b_2/2$	$b_2/1$	$b_3/1$

(б)

δ/λ	b_1	b_2	b_3	b_4
x_1	$b_4/2$	$b_3/2$	$b_3/2$	$b_3/1$
x_2	$b_1/2$	$b_1/1$	$b_1/2$	$b_1/1$
x_3	$b_2/1$	$b_1/2$	$b_3/1$	$b_3/2$

10. (а) Побудуємо мінімальні автомати Z_1 і Z_2 для A_1 і A_2: для автомата A_1 — $b_1 = \{a_1, a_5, a_8\}$, $b_2 = \{a_2, a_9\}$, $b_3 = \{a_3\}$, $b_4 = \{a_4, a_6, a_7\}$, для A_2 — $c_1 = \{a_1, a_3, a_4\}$, $c_2 = \{a_2, a_5\}$, $c_3 = \{a_6\}$, $c_4 = \{a_7\}$. Суміщені таблиці мінімальних автоматів Z_1 і Z_2 мають такий вигляд:

Z_1

δ_1/λ_1	b_1	b_2	b_3	b_4
x_1	$b_3/1$	$b_3/2$	$b_3/2$	$b_1/2$
x_2	$b_4/1$	$b_4/1$	$b_4/2$	$b_4/2$
x_3	$b_3/2$	$b_4/2$	$b_3/1$	$b_2/1$

Z_2

δ_2/λ_2	C_1	c_2	c_3	c_4
x_1	$C_4/2$	$c_3/2$	$c_3/2$	$c_3/1$
x_2	$C_1/2$	$c_1/1$	$c_1/2$	$c_1/1$
x_3	$C_2/1$	$c_1/2$	$c_3/1$	$c_2/2$

Гомоморфізм автомата A_1 на автомат Z_1: $\gamma_1 = \{(a_1, b_1), (a_5, b_1), (a_8, b_1), (a_2, b_2), (a_9, b_2), (a_3, b_3), (a_4, b_4), (a_6, b_4), (a_7, b_4)\}$, гомоморфізм автомата A_2 на автомат Z_2: $\gamma_2 = \{(a_1, c_1), (a_3, c_1), (a_4, c_1), (a_2, c_2), (a_5, c_2), (a_6, c_3), (a_7, c_4)\}$. Відображення $\eta = \{(b_1, c_4), (b_2, c_2), (b_3, c_3), (b_4, c_1)\}$ є ізоморфізмом автомата Z_1 на автомат Z_2. Отже, автомати A_1 і A_2 невідрізнювані. (б) Побудуємо мінімальні автомати Z_1 і Z_2 для A_1 і A_2: для автомата A_1 — $b_1 = \{a_1, a_4, a_5, a_6\}$, $b_2 = \{a_2, a_3, a_9\}$, $b_3 = \{a_7, a_8\}$, для автомата A_2 — $c_1 = \{a_1\}$, $c_2 = \{a_2, a_3, a_5, a_6\}$, $c_3 = \{a_4, a_7\}$.

Z_1

δ_1/λ_1	b_1	b_2	b_3
x_1	$b_2/2$	$b_3/1$	$b_2/2$
x_2	$b_1/1$	$b_3/2$	$b_3/1$
x_3	$b_2/1$	$b_1/2$	$b_1/1$

Z_2

δ_2/λ_2	c_1	C_2	c_3
x_1	$c_3/2$	$C_3/2$	$c_2/1$
x_2	$c_1/1$	$C_2/1$	$c_2/2$
x_3	$c_3/1$	$C_1/1$	$c_1/2$

Принцип побудови гомоморфізмів γ_1 і γ_2 див. у попередній задачі. Відображення $\eta = \{(b_1, c_1), (b_2, c_3), (b_3, c_2)\}$ є ізоморфізмом Z_1 на Z_2, тобто автомати A_1 і A_2 невідрізнювані.

7.6. 4. (а) Результатом мінімізації автомата A_1 є автомат Z_1, у якому $b_1 = \{a_1, a_3, a_5, a_8\}$, $b_2 = \{a_2, a_9\}$, $b_3 = \{a_4, a_6, a_7\}$, а результатом мінімізації автомата A_2 є автомат Z_2, у якому $c_1 = \{a_1, a_3, a_4\}$, $c_2 = \{a_2, a_5\}$, $c_3 = \{a_6, a_7\}$. Відмічені таблиці переходів автоматів Z_1 і Z_2 мають такий вигляд:

Z_1

μ_1	1	2	1
δ_1	b_1	b_2	b_3
x_1	b_1	b_1	b_1
x_2	b_3	b_3	b_3
x_3	b_1	b_3	b_2

Z_2

μ_2	1	2	1
δ_2	c_1	c_2	c_3
x_1	c_3	c_3	c_3
x_2	c_1	c_1	c_1
x_3	c_2	c_1	c_3

Гомоморфізм автомата A_1 на автомат Z_1: $\gamma_1 = \{(a_1, b_1), (a_3, b_1), (a_5, b_1), (a_8, b_1), (a_2, b_2), (a_9, b_2), (a_4, b_3), (a_6, b_3), (a_7, b_3)\}$, гомоморфізм автомата A_2 на автомат Z_2: $\gamma_2 = \{(a_1, c_1), (a_3, c_1), (a_4, c_1), (a_2, c_2), (a_5, c_2), (a_6, c_3), (a_7, c_3)\}$. Відображення $\eta = \{(b_1, c_3), (b_2, c_2), (b_3, c_1)\}$ буде ізоморфізмом автомата Z_1 на автомат Z_2. Отже, автомати A_1 і A_2 невідрізнювані. (б). Мінімізуючи автомат A_1, отримаємо автомат Z_1, у якому $b_1 = \{a_1, a_4, a_5, a_6\}$, $b_2 = \{a_2, a_3, a_9\}$, $b_3 = \{a_7, a_8\}$, а мінімізуючи автомат A_2, отримаємо Z_2, у якому $c_1 = \{a_1\}$, $c_2 = \{a_2, a_3, a_5, a_6\}$, $c_3 = \{a_4, a_7\}$. Відмічені таблиці переходів автоматів Z_1 і Z_2 мають такий вигляд:

Z_1

μ_1	1	1	2
δ_1	b_1	b_2	b_3
x_1	b_2	b_3	b_2
x_2	b_1	b_3	b_3
x_3	b_2	b_1	b_1

Z_2

μ_2	1	2	1
δ_2	c_1	c_2	c_3
x_1	c_3	c_3	c_2
x_2	c_1	c_2	c_2
x_3	c_3	c_1	c_1

Принцип побудови гомоморфізмів γ_1 і γ_2 див. у попередній задачі. Відображення $\eta = \{(b_1, c_1), (b_2, c_3), (b_3, c_2)\}$ є ізоморфізмом Z_1 на Z_2, тобто автомати A_1 і A_2 невідрізнювані. **5.** Відмічені таблиці переходів шуканих автоматів Мура мають такий вигляд:

(а)

μ	–	–	–	1	2	2	1	1	1
δ	b_{10}	b_{20}	b_{30}	b_{11}	b_{12}	b_{21}	b_{22}	b_{31}	b_{32}
x_1	b_{11}	b_{21}	b_{31}	b_{21}	b_{11}	b_{21}	b_{31}	b_{11}	b_{11}
x_2	b_{12}	b_{22}	b_{32}	b_{22}	b_{12}	b_{22}	b_{22}	b_{12}	b_{12}

(б)

μ	–	–	–	2	1	1	1	1	2
δ	b_{10}	b_{20}	b_{30}	b_{11}	b_{12}	b_{21}	b_{22}	b_{31}	b_{32}
x_1	b_{11}	b_{21}	b_{31}	b_{31}	b_{11}	b_{21}	b_{11}	b_{31}	b_{21}
x_2	b_{12}	b_{22}	b_{32}	b_{32}	b_{12}	b_{22}	b_{12}	b_{32}	b_{22}

7.7. 1. (а) Автомат Мілі, що реалізує задане відображення, перебуває в стані 1, якщо двома останніми вхідними сигналами були 11, у стані 2 — якщо 10, у стані 3 — якщо 00 і в стані 4 — якщо 01. Його суміщена таблиця має такий вигляд:

δ/λ	1	2	3	4
0	2/1	3/1	3/0	2/0
1	1/1	4/1	4/0	1/0

(б) Суміщені таблиці шуканого автомата для $l = 0$ та для $l > 0$ мають такий вигляд:

$l = 0$

δ/λ	1	2
0	1/1	2/0
1	2/0	2/0

$l > 0$

δ/λ	1	2	3	…	l	$l+1$	$l+2$	$l+3$
0	1/0	2/0	3/0	…	l/0	$l+2$/1	$l+2$/1	$l+3$/0
1	2/0	3/0	4/0	…	$l+1$/1	$l+3$/0	$l+3$/0	$l+3$/0

(в), (г) Нижче подано суміщені таблиці відповідних автоматів. Автомат (в) перебуває в стані 0, якщо поточна кількість одиниць у вхідному слові дорівнює $3k$, у стані 1 — якщо вона дорівнює $3k + 1$, у стані 2 — якщо $3k + 2$, $k = 0, 1, 2, \ldots$. Автомат (г) перебуває в стані 1, якщо на його вхід подано сигнал $x(2t - 1)$, у стані 2 — якщо подано сигнал $x(2t)$ і попереднім вихідним сигналом був 0, у стані 3 — якщо подано сигнал $x(2t)$ і попереднім вихідним сигналом була 1.

(в)

δ/λ	0	1	2
0	0/1	1/0	2/0
1	1/0	2/0	0/1

(г)

δ/λ	1	2	3
0	2/0	1/0	1/1
1	3/1	1/1	1/0

(д), (е) Нижче подано суміщені таблиці відповідних автоматів. Автомат (д) перебуває в стані 1 у початковий момент часу, у стані 2 — якщо попереднім вхідним сигналом був 0, у стані 3 — якщо попереднім вхідним сигналом була 1. Автомат (е) перебуває в стані 1, якщо на його вхід подано сигнал $x(3t - 2)$, у стані 2 — якщо подано сигнал $x(3t - 2)$ і попереднім вхідним сигналом був 0, у стані 3 — якщо подано сигнал $x(3t - 2)$ і попереднім сигналом була 1, у стані 4 — якщо подано сигнал $x(3t)$ і попереднім сигналом був 0, у стані 5 — якщо подано сигнал $x(3t)$ і попереднім сигналом була 1, $t = 1, 2, \ldots$.

(д)

δ/λ	1	2	3
0	2/1	2/0	2/1
1	3/1	3/1	3/0

(е)

δ/λ	1	2	3	4	5
0	2/1	4/0	5/1	1/0	1/0
1	3/0	4/0	5/1	1/0	1/1

2. Нижче подано суміщену таблицю шуканого автомата. Автомат перебуває в стані 0, якщо на його вхід подано сигнал $x(4t)$, у стані 1 — якщо $x(4t + 1)$, у стані 2 — якщо $x(4t + 2)$, у стані 3 — якщо $x(4t + 3)$, $t = 0, 1, 2, \ldots$.

δ/λ	0	1	2	3
0	1/1	2/0	3/0	0/0
1	1/0	2/1	3/1	0/1

3. (а) Позначимо $a_{i_k} = \delta(a_1, x^k)$. Тоді внаслідок скінченності множини U для послідовності станів $a_{i_0}, a_{i_1}, a_{i_2}, \ldots, a_{i_s}$ існуватимуть такі значення j та l, що $i_j = i_l$, $0 \le j < l \le s$. Нехай j і l — найменші числа, для яких виконуються ці умови. Тоді $\varphi_A^{i_j}(x^{l-j})$ є періодом вихідної послідовності z, а $\varphi_A^1(x^j)$ — її передперіодом. *Зауваження.* Якщо автомат A не мінімальний і вага його автоматного відображення φ_A^1 дорівнює p, то всі наведені вище міркування матимуть місце для послідовності станів $a_{i_0}, a_{i_1}, a_{i_2}, \ldots, a_{i_p}$, $p < s$. (б), (в) Із доведення пункту (а) випливає, що довжина передперіоду послідовності z не перевищує $p - 1$ ($j = p - 1$, $l = p$), а максимальна довжина періоду дорівнює p ($j = 0$, $l = p$). Для зведеного автомата $p = s$. **4.** Аналізуючи розв'язок задачі 7.7.3(а), дістанемо, що сума довжин періоду і передперіоду відповідної вихідної послідовності дорівнює $|x^{l-j}| + |x^j| = (l - j) + j = l$, і для l викону-ється $l \le p$. **5.** Нехай на вхід автомата-перетворювача $A = (X, Y, U, \delta, \lambda)$ подано періодичну послідовність вхідних сигналів $vwww\ldots w\ldots$, $v, w \in X^*$. Припустимо, що вага автоматного відображення φ_A^1 дорівнює p. Розглянемо послідовність станів $a_{i_t} = \delta(a_1, vw^t)$ і відповідну послідовність залишкових відображень $\varphi_A^{i_t}$ для $t = 0, 1, 2, \ldots, p$. З означення ваги випливає, що для φ_A^1 існуватимуть такі значення j і l, що залишкові відображення $\varphi_A^{i_j}$ і $\varphi_A^{i_l}$ збігаються або ж, що те саме — стани a_{i_j} і a_{i_l} невідрізнювані, $0 \le j < l \le p$. Тоді за формулою (7.4) отримаємо $q = \varphi_A^1(vwww\ldots w\ldots) = \varphi_A^1(vw^j)\varphi_A^{i_j}(w^{l-j})\varphi_A^{i_{l_1}}(w^{l-j})\varphi_A^{i_{l_2}}(w^{l-j})\ldots$, де $a_{i_l} = \delta(a_{i_j}, w^{l-j})$, $a_{i_{l_1}} = \delta(a_{i_l}, w^{l-j})$ і $a_{i_{l_r}} = \delta(a_{i_{l_{r-1}}}, w^{l-j})$, $r = 2, 3, \ldots$. Оскільки стани a_{i_j} і a_{i_l} невідрізнювані, то невідрізнюваними будуть і відповідні стани a_{i_l} і $a_{i_{l_1}}$, у які автомат A переходить під дією того самого слова w^{l-j} (див. задачу 7.5.7). Подібні міркування обґрунтовують невідрізнюваність станів $a_{i_{l_1}}$ і $a_{i_{l_2}}$, $a_{i_{l_2}}$ та $a_{i_{l_3}}$ і т. д. Отже, відповідна вихідна послідовність q матиме вигляд $uzz\ldots z\ldots$, де $u = \varphi_A^1(vw^j)$ і $z = \varphi_A^{i_j}(w^{l-j})$. **6.** Аналізуючи розв'язання попередньої задачі, дістанемо, що довжина періоду послідовності q дорівнює $|z| = |w^{l-j}| = k(l - j)$, а довжина передперіоду — $|u| = |vw^j| = d + kj$, бо за умовою $|w| = k$ і $|v| = d$. Оскільки для j та l виконуються нерівності $0 \le j < l \le p$, то максимальним значенням для різниці $l - j$ є p ($j = 0$ й $l = p$), а максимальним значенням для j є $p - 1$. Отже, максимальним значенням для довжини періоду послідовності q є pk, а для довжини передперіоду — $d + k(p - 1)$.

7.8. 1. Таблиці переходів відповідних автоматів мають такий вигляд. Для зручності в них замість стану a_i записуємо тільки його індекс i. Заключні стани підкреслено.

(а)

δ	1	2	3	4	5	6	**7**	8	9	**10**	f
x_1	2	3	4	6	f	f	9	**10**	**10**	f	f
x_2	f	f	5	**7**	8	**10**	F	f	f	f	f

(б)

δ	**1**	2	3	4	5	6	7	**8**	9	**10**	11	f
x_1	2	3	f	9	6	7	f	f	**10**	f	f	f
x_2	f	4	5	f	f	f	**8**	f	f	11	**8**	f

2. Автомат A' відрізняється від автомата A тільки тим, що множина його заключних станів $F' = F \cup \{a_1\}$, де F — множина заключних станів автомата A. **3.** Таблиця переходів шуканого автомата має такий вигляд. Заключні стани підкреслено.

δ	1	2	3	4	**5**	6	7	8	**9**	10
0	2	3	4	**5**	**5**	10	7	7	7	10
1	6	10	3	3	3	7	8	**9**	**9**	10

4. З означення задачі 7.2.6 випливає, що елемент $m_{ij}(w)$ матриці $M(w)$ дорівнює 1, якщо $a_j = \delta(a_i, w)$, і дорівнює 0 — в іншому разі. Тоді координата t_i вектора $pM(w)$ дорівнює 1, якщо $a_i = \delta(a_1, w)$, тобто якщо слово w переводить автомат A зі стану a_1 у стан a_i, інакше $t_i = 0$, $i = 1, 2, \ldots, s$. Отже, $pM(w)f = 1$ тоді й тільки тоді, коли слово w переводить автомат A з початкового стану a_1 у деякий заключний стан. Рівність $pM(w)f = 1$ в цьому разі рівносильна нерівності $pM(w)f \neq 0$ або $pM(w)f > 0$. **5.** Припустімо, що існує скінченний автомат A, який розпізнає деяку неперіодичну послідовність вхідних символів w і тільки її. Розглянемо послідовність станів $a_{i_k} = \delta(a_1, \eta_k(w))$, $k = 0, 1, 2, \ldots$. З умови випливає, що всі стани a_{i_k} заключні. Оскільки автомат A скінченний, то існуватимуть такі значення j і l, що $i_j = i_l$, $j < l$. Тоді автомат A розпізнає також періодичну послідовність вхідних символів $w' = vzz\ldots z\ldots$, де $v = x_{i_1}x_{i_2}\ldots x_{i_j}$ і $z = x_{i_{j+1}}x_{i_{j+2}}\ldots x_{i_l}$. Цей висновок суперечить зробленим припущенням. **6.** Доведення проведемо методом від супротивного. Припустімо, що подія P є зображуваною в деякому скінченному автоматі $A = (X, U, \delta, a_1, F)$. Розглянемо послідовність станів $a_{i_k} = \delta(a_1, x_1^{k^2})$, $k = 0, 1, 2, \ldots$; $x_1 \in X$. Зі скінченності множини станів U випливає, що існуватимуть такі числа j і l, що $i_j = i_l$, $j < l$. За умовою слова $x_1^{l^2}$ і $x_1^{(l+1)^2}$ розпізнаються, а з рівності станів $a_{i_j} = \delta(a_1, x_1^{j^2})$ і $a_{i_l} = \delta(a_1, x_1^{l^2})$ випливає, що й слово $x_1^{(l^2+2j+1)}$ також розпізнаються автоматом A. Отже, $l^2 + 2j + 1$ є квадратом якогось натурального числа t. Із нерівностей $l^2 < l^2 + 2j + 1 < (l + 1)^2$ або $l^2 < t^2 < (l + 1)^2$ випливатиме хибний висновок, що існує натуральне число t, для якого $l < t < l + 1$, $l \in N$. Отримана суперечність спростовує зроблене вище припущення. **7.** (а) Позначимо через x_1 вхідний сигнал, що відповідає букві, x_2 — вхідний сигнал, що відповідає цифрі, і через x_3 — вхідний сигнал, що відповідає будь-якому іншому символу вхідного алфавіту. Таблицю переходів відповідного автомата наведено нижче (заключний стан автомата — 2). (б) Позначимо через Ц вхідний сигнал, що відповідає цифрі, через З — сигнал, що відповідає знаку (+ або –), через К — сигнал, що відповідає крапці (комі) між цілою та дробовою частинами числа, і через І — сигнал, що відповідає будь-якому іншому символу вхідного алфавіту. Таблиця переходів автомата має такий вигляд (заключні стани — 4 та 5).

(а)

δ	1	2	3
x_1	2	2	3
x_2	3	2	3
x_3	3	3	3

(б)

δ	1	2	3	4	5	6
Ц	4	4	5	4	5	6
З	2	6	6	6	6	6
К	3	3	6	5	6	6
І	6	6	6	6	6	6

(в) Простим арифметичним виразом вважатимемо вираз, у якому операндами є ідентифікатори або константи, операціями — деякі бінарні операції, а вкладеність дужок — не більше 2. Позначимо через x_1 операнд, x_2 — ліву дужку (, x_3 — праву дужку) і через x_4 — знак бінарної операції. Таблиця переходів автомата має такий вигляд (заключними станами є 4 та 9).

δ	1	2	3	4	5	6	7	8	9	10	11	12	13	14
x_1	2	14	4	14	6	14	8	14	14	11	14	13	14	14
x_2	5	14	5	14	10	14	11	14	14	14	14	14	14	14
x_3	14	14	14	14	14	14	14	9	14	14	14	14	14	14
x_4	14	3	14	3	14	7	14	7	3	14	12	14	12	14

(г) Коментарем вважатимемо послідовність символів, що має такий вигляд: /* < текст коментаря > */. Позначимо через x_1 символ /, через x_2 — символ * і через x_3 — довільний символ, відмінний від / і *. Таблиця автомата має такий вигляд (заключним станом є стан 5).

δ	1	2	3	4	5	6
x_1	2	6	3	5	3	6
x_2	6	3	4	4	4	6
x_3	6	6	3	3	3	6

7.9. 1. Нехай $w \in PQ \Rightarrow w = w_1w_2$, $w_1 \in P$, $w_2 \in Q \Rightarrow w = w_1w_2$, $w_1 \in R$, $w_2 \in S \Rightarrow w \in RS$. Обернене твердження не є справедливим. Наприклад, для $P = \{10\}$, $Q = \{010\}$, $R = \{1, 01\}$, $S = \{0010, 11\}$, $X = \{0, 1\}$. **2.** $P = \{0\}$, $Q = \{1, 00, 101, 111, 1000, 10111\}$. **3.** (а) Якщо $w \in P \cup Q$, то $w \in P$ або $w \in Q$. Кожен елемент v, що належить одночасно P і Q, у лівій частині $|P \cup Q|$ підраховується один раз, а в правій частині $|P| + |Q|$ — двічі. Отже, кількість елементів у $P \cup Q$ не перевищує кількості елементів у P і в Q. Рівність виконується тоді й тільки тоді, коли $P \cap Q = \varnothing$. (б) Одне й те саме слово $w \in PQ$ можна утворити, взагалі кажучи, з більше ніж однієї пари слів, що належать P та Q. Наприклад, $0110 = w_1w_2$ для $w_1 = 0$ і $w_2 = 110$, $w_1 = 01$ і $w_2 = 10$ та ін. Тому в множині PQ слів не більше ніж $|P||Q|$. Рівність виконується тоді й тільки тоді, коли кожне слово з PQ можна утворити не більше ніж одним способом зі слів подій P і Q. **4.** (а) $\{e, ab, b, abab, abb, bab, bb, ababb, abbab, babab, abbb, babb, bbab, bbb, abbbb, babbb, bbabb, bbbab, bbbb, bbbbb\}$. **5.** (а) 8. (б) 29. (в) 2^n. **6.** (а), (б) $P = \{011, 101\}$. Не існує. Якщо $Q = \{w_1, w_2, \dots, w_n\}$ — множина всіх шестибуквених слів із P^*, то множина всіх дванадцятибуквених слів події P^* містить у собі множину $\{w_iw_j \mid i, j = 1, 2, \dots, n\}$, кількість елементів у якій не менша від $|Q|$. **7.** Якщо $w \in P^*$, то слово w є результатом конкатенації деяких слів із P. Оскільки $P \subseteq Q$, то $w \in Q^*$. **8.** Оскільки $P \subseteq Q$, то $P^* \subseteq Q^*$ (див. попередню задачу). Єдине слово, яким відрізняються події P та Q, — це слово $aaaa$. Однак воно може бути утворене в результаті конкатенації слів aa і aa, що належать P. Тому всі слова з Q^* можна утворити за допомогою конкатенації зі слів події P. Отже, $Q^* \subseteq P^*$. **9.** $P \subseteq Q$, тому $P^* \subseteq Q^*$. Водночас $aaa \in Q^*$, однак слово aaa не можна утворити за допомогою конкатенації зі слів події P, тобто $aaa \notin P^*$. **10.** (а) Див. 7.9.8. (б) $P = \{aa, ba, baba\}$, $Q = \{aa, ba, aaaa\}$, далі див. 7.9.8. **11.** (а), (б) Нехай $|X| > 1$ і $a \in X$. Тоді обидва співвідношення виконуються для $P = \{a^k \mid k = 0, 1, 2, \dots\}$. **12.** (а) Неважко переконатися, що в будь-якій події P є неподільне слово w, тобто слово, для якого не існує слів $s, t \in P$, відмінних від порожнього слова e і таких, що $w = st$. Нехай w — неподільне слово даної події P. Тоді згідно з умовою $w \in P^2$. Останнє співвідношення можливе лише тоді, коли $e \in P$. (б) Методом математичної індукції можна довести, що $P^n = P$ для всіх $n \in N$ (використати результат задачі 7.9.1). Тоді матимемо $P^* = \{e\} \cup P \cup P^2 \cup P^3 \cup \dots = \{e\} \cup P \cup P \cup P \cup \dots = \{e\} \cup P$. Враховуючи твердження попереднього пункту, отримаємо рівність $P^* = P$. **13.** Покладемо $P = \{a\}$, $Q = \{b\}$ для алфавіту $X = \{a, b\}$. Тоді $ab \in (P \cup Q)^*$, однак $ab \notin P^* \cup Q^*$. Необхідною й достатньою умовою виконання даної рівності є $P \subseteq Q$ або $Q \subseteq P$. Справді, нехай має місце рівність $(P \cup Q)^* = P^* \cup Q^*$, однак не виконується жодне з включень, тобто існує $a \in P$ і $a \notin Q$ та $b \in Q$ і $b \notin P$. Тоді $ab \in (P \cup Q)^*$, але $ab \notin P^* \cup Q^*$, бо $ab \notin P^*$ і $ab \notin Q^*$. Навпаки, нехай $P \subseteq Q$. Тоді $P \cup Q = Q$, $P^* \subseteq Q^*$ і $P^* \cup Q^* = Q^* = (P \cup Q)^*$. Подібні міркування можна провести й у разі, коли $Q \subseteq P$. **14.** $P = \{e, aaa, aab, aba, abb, baa, bab, bba, bbb\}$. **15.** Нехай $p, q \in P^*$. Тоді $p \in P^k$ і $q \in P^m$ для деяких цілих невід'ємних k і m. Звідси $pq \in P^{k+m}$, тому $pq \in P^*$. **16.** Нехай Q — подія, замкнена відносно конкатенації, тоді $Q^2 \subseteq Q$. Звідси $Q^k \subseteq Q$ для всіх $k \in N$, отже $Q^+ = Q$. З умови $P \subseteq Q$ випливає $P^+ \subseteq Q^+$ (див. 7.9.7). Отже, $P^+ \subseteq Q$. **17.** Безпосередньо з означення замикання випливає, що $[P] = P \cup P^2 \cup P^3 \cup \dots$. Отже, $[P] = P^+$. **18.** Найменшою множиною, що містить P і замкнена відносно конкатенації, є $[P]$ (див. означення замикання множини в розд. 2.5). Тоді за результатом попередньої задачі $P^+ = [P]$. **19.** $(P \cup Q)^+$. **20.** (а) $1^* \cup (1^*01^*01^*01^*)^*$. (б) $0^*10^*10^*10^*$. (в) $\{0,1\}^*(00 \cup 11)$. (г) $(01)^*00(10)^* \cup (10)^*11(01)^*$. (д) $(0 \cup 01 \cup 011)^* \cup (0 \cup 10 \cup 110)^* \cup (0 \cup 01 \cup 101)^* \cup (0 \cup 10 \cup 101)^*$. (е) $(0^*10^*10^*)^*$. (є) $X^6 0 X^*$. (ж) $X^* 1 X^2$.

(з) $P\cup 1P\cup P0\cup 1P0$, де $P=(01)^*$. (и) $(1^*(00)^*1^*)^*$. **21.** 3. $w\in P(Q\cup T)\Leftrightarrow$ $\Leftrightarrow w=w_1w_2$, $w_1\in P$, $w_2\in Q\cup T\Leftrightarrow w=w_1w_2$, $w_1\in P$, ($w_2\in Q$ або $w_2\in T$) $\Leftrightarrow$ $\Leftrightarrow w=w_1w_1$, $((w_1\in P,\ w_2\in Q)$ або $(w_1\in P,\ w_2\in T))\Leftrightarrow w\in PQ$ або $w\in PT\Leftrightarrow$ $\Leftrightarrow w\in PQ\cup PT$. 5. Якщо $w\in(P^*)^*$, то w є результатом конкатенації деяких слів $w_1, w_2, \ldots, w_k\in P^*$. У свою чергу, кожне слово w_i, $i=1, 2, \ldots, k$ є результатом конкатенації слів із P. Отже, $w\in P^*$. Обернене включення випливає з означення операції ітерації. 7. $PP^*=P^+$, тому $\{e\}\cup PP^*=P^*$. 9. Використовуючи означення ітерації та властивості 1–4 операції об'єднання і конкатенації, кожен із даних виразів можна перетворити до вигляду $\cup P^{k_1}Q^{m_1}P^{k_2}Q^{m_2}\ldots P^{k_n}Q^{m_n}$, де об'єднання здійснюється за всіма можливими наборами $k_1, m_1, k_2, m_2, \ldots, k_n, m_n$ невід'ємних цілих чисел, $n=1, 2, \ldots$. **22.** Дана подія складається з усіх слів в алфавіті $\{a, b\}$: (а) що починаються з a; (б) що мають парну довжину; (в) у яких різниця між кількостями входжень символів a і b є парною; (г) які закінчуються або на a, або на $abbb$; (д) які утворюються з a і bb та починаються з a; (е) які містять однакову кількість фрагментів a^k і b^l, де k, l — непарні числа, а також із порожнього слова e; (є) які утворюються із фрагментів a^k та b^l для непарних k і l, та з порожнього слова e. Непорожні слова починаються з фрагмента b^m, за яким слідує однакова кількість a^k і b^l; (ж) утворених з aa і ba, та порожнього слова e. **23.** Множина станів $U=\{1, 2, 3, 4, 5\}$, 1 — початковий стан, 5 — заключний стан. Функцію переходів задамо такими співвідношеннями: $\delta(1, a)=\delta(2, a)=\delta(3, a)=2$, $\delta(1, b)=\delta(4, b)=1$, $\delta(2, b)=3$, $\delta(3, b)=4$, $\delta(4, a)=\delta(5, a)=\delta(5, b)=5$. **24.** Позначимо подію, яку задає регулярний вираз у лівій частині рівності, через P_1, а подію правої частини — через P_2. (а) P_2 — це множина всіх слів в алфавіті $\{a, b\}$, тому $P_1\subseteq P_2$. Нехай $w\in P_2$. Розглянемо два випадки: слово w містить підслово ab і протилежну ситуацію — слово w не містить підслово ab. У першому випадку $w=w_1abw_2$, де $w_1, w_2\in\{a, b\}^*=P_2$, а в другому — $w=b^ka^l$, $k, l=0, 1, 2, \ldots$, тобто $w\in\{b\}^*\{a\}^*$. Отже, $P_2\subseteq P_1$, тому $P_1=P_2$. (б) $w\in P_1\Leftrightarrow w=(ab)^ka$, $k=0, 1, 2, \ldots$, $\Leftrightarrow w=a(ba)^k$, $k=0, 1, 2, \ldots\Leftrightarrow w\in P_2$. (в) $P_1\subseteq P_2$ (див. (а)). З іншого боку, оскільки $\{a, b\}\subseteq\{a\}^*\cup\{b\}$, то $P_2\subseteq P_1$ (див. 7.9.7). (г) $w\in P_1\Leftrightarrow w=(a^{k_1}b)(a^{k_2}b)\ldots(a^{k_m}b)a^n$, $k_1, k_2, \ldots, k_m$, $n=0, 1, 2, \ldots\Leftrightarrow$ $\Leftrightarrow w=a^{k_1}(ba^{k_2})(ba^{k_3})\ldots(ba^n)$, $k_1, k_2, \ldots, k_m$, $n=0, 1, 2, \ldots\Leftrightarrow w\in P_2$. (д) $P_2\subseteq P_1$, бо $e\in P_1$ і $\{a, bb\}^*aa\subseteq\{\{a, bb\}^*aa\}^*$. Нехай $w\in P_1$. Тоді $w=e$ або $w=w_1aa$, де w_1 — слово, що складається з a і bb. Аналізуючи регулярний вираз у

правій частині, доходимо висновку, що $w\in P_2$. (е) Будь-яке слово з P_1 і P_2 має вигляд a^k, $k=0, 1, 2, \ldots$. (є) Див. (г). (ж) $w\in P_1\Leftrightarrow w=aa^{k_1}(ba)^{m_1}a^{k_2}(ba)^{m_2}\ldots a^{k_n}(ba)^{m_n}b$, $k_i, m_i=0, 1, 2, \ldots$; $i=1, 2, \ldots, n\Leftrightarrow w=aa^{k_1}b(a(ba)^{m_1-1}a^{k_2}(ba)^{m_2}\ldots a^{k_n}(ba)^{m_n}b)$, $k_i, m_i=0, 1, 2, \ldots$; $i=1, 2, \ldots, n\Leftrightarrow w=aa^{k_1}b(ab)^{m_1-1}aa^{k_2}b(ab)^{m_2-1}aa^{k_3}b\ldots aa^{k_n}b(ab)^{m_n-1}$, $k_i\geq 0$, $m_i\geq 1$, $i=1, 2, \ldots, n\Leftrightarrow w\in P_2$, бо $aa^{k_i}b\in\{a\}\{a\}^*\{b\}$ і $(ab)^{m_i-1}$, $aa^{k_i}b\in(\{a\}\{a\}^*\{b\})^*$, $i=1, 2, \ldots, n$. **25.** Позначимо подію, яку задає вираз у лівій частині, через P_1, а подію з правої частини — через P_2. (а) $P_1=P_2$, бо $w\in P_1\Leftrightarrow w=(ab)^{k_1}a^{m_1}(ab)^{k_2}a^{m_2}\ldots a^{k_n}(ab)^{m_n}a$, $k_i, m_i=0, 1, 2, \ldots$; $i=1, 2, \ldots, n\Leftrightarrow w=a(ba)^{k_1}a^{m_1}(ba)^{k_2}a^{m_2}\ldots(ba)^{m_n}$, $k_i, m_i=0, 1, 2, \ldots$; $i=1, 2, \ldots, n\Leftrightarrow w\in P_2$. (б) $P_1\neq P_2$, бо $e\notin P_1$ і $e\in P_2$. (в) Див. (б). (г) $P_1\neq P_2$, бо $ba\in P_1$ і $ba\notin P_2$. (д) $P_2\subseteq P_1$. Навпаки, нехай $w\in P_1$, тоді $w=a^{k_1}b^{m_1}a^{k_2}b^{m_2}\ldots a^{k_n}b^{m_n}$, $k_i, m_i=0, 1, 2, \ldots$; $i=1, 2, \ldots, n$. Неважко переконатися, що $w\in\{a\}^*(\{b\}\{a\}^*)^*$, тобто $w\in P_2$. (е) P_1 — це подія, що складається з усіх слів у алфавіті $\{a, b\}$, які починаються з a або з b, а також порожнього слова e. Отже, $P_1=P_2$.

7.10. 2. (а) Джерело складається з початкової і заключної вершин S і F та десяти вершин $v_0, v_1, \ldots, v_9$. У кожну з вершин v_i з початкової вершини S веде порожня дуга. Із вершини v_i у вершину v_i веде 9 дуг (петель), позначених цифрами з множини $\{0, 1, \ldots, 9\}\backslash\{i\}$, а з v_i у F веде дуга з відміткою i. (б), (в). Таблиці джерел зображено нижче. Початковими вершинами джерел є 1. Заключними вершинами для джерела (б) є 7, а для джерела (в) — 3, 4, …, 9.

(б)

δ	1	2	3	4	5	6	7
0	2	7	–	–	–	7	–
1	1	3	4	5	6	3	7

(в)

δ	1	2	3	4	5	6	7	8	9
0	2	2	4	5	6	7	8	9	2
1	3	3	3	3	3	3	3	3	3

6. (а) Шукане джерело D побудуємо, використовуючи граф G автомата A, що зображує мову L: змінимо напрямок усіх дуг у G на протилежний, додамо до G нову вершину v_0, яку вважатимемо початковою вершиною D, і з'єднаємо її порожніми дугами з усіма заключними станами A. Заключною вершиною D є початковий стан A. (б) Див. (а).

7.11. 1. (а) Множина станів $U = \{1, 2, 3, 4, 5, 6\}$. Початковий стан — 1, заключні — 2, 4, 5. Функцію переходів задамо такими співвідношеннями: $\delta(1, a) = \delta(2, b) = 2$, $\delta(1, b) = \delta(4, b) = 3$, $\delta(2, a) = \delta(4, a) = \delta(5, b) = \delta(6, a) = \delta(6, b) = 6$, $\delta(3, a) = 4$, $\delta(3, b) = \delta(5, a) = 5$. **3.** До джерела D слід додати порожню дугу, що веде з початкової вершини у заключну. **4.** До джерела D додати нову вершину w і провести дугу 0 із початкової вершини у w та дугу 1 із w у заключну вершину. **5.** Множиною вершин джерела D є $\{1, 2, \ldots, n + 1\}$, вершина 1 — початкова, а вершина $n + 1$ — заключна. Дуги джерела D задамо за допомогою таких співвідношень: $\delta(1, 0) = 1$, $\delta(1, 1) = \{1, 2\}$ і $\delta(i, 0) = \delta(i, 1) = i + 1$ для $i = 2, 3, \ldots, n$. Для формального обгрунтування того, що D — шукане джерело, слід довести індукцією за довжиною вхідного слова w такі два твердження: 1) шлях, що відповідає довільному вхідному слову p, веде в джерелі D з вершини 1 у вершину 1; 2) шлях веде у джерелі D з початкової вершини 1 у вершину $i + 1$ тільки тоді, коли йому відповідає вхідне слово w, в якому i-й символ від кінця є 1. **6.** Детерміноване джерело (скінченний автомат) A, що зображує подію P, має за допомогою своїх вершин (станів) "запам'ятовувати" всі 2^n можливих комбінацій останніх n символів вхідного слова. Отже, воно міститиме щонайменше 2^n відповідних вершин. В іншому разі існуватиме стан a, в який A потрапляє після появи на його вході двох різних слів $w_1 = x_1x_2\ldots x_n$ і $w_2 = y_1y_2\ldots y_n$ в алфавіті $\{0, 1\}$. Оскільки $w_1 \neq w_2$, то існує такий індекс k, що $x_k \neq y_k$, $k = 1, 2, \ldots, n$. Припустимо, що $x_k = 1$ та $y_k = 0$. Якщо $k = 1$, то $w_1 \in P$ і $w_2 \notin P$. Отже, стан a має бути одночасно заключним і незаключним. Якщо ж $k > 1$, то розглянемо стан b, у який автомат A потрапляє зі стану a після подання слова 0^{k-1}. Знову отримаємо суперечність, бо стан b, у який ведуть в A слова $w_1 0^{k-1}$ і $w_2 0^{k-1}$ із початкового стану, має бути й не бути заключним, тому що $w_1 0^{k-1} \in P$ і $w_2 0^{k-1} \notin P$. **7.** Шуканий скінченний автомат можна отримати в результаті детермінізації джерела, що зображує подію P^{-1} (див. 7.10.6 (а)). **8.** 11.

7.12. 1. (а) $e \cup x_1(x_2 \cup x_1x_1)^*x_1 \cup \left(x_2 \cup x_1(x_2 \cup x_1x_1)^*x_1x_2\right)\left(x_1 \cup x_2(x_2 \cup x_1x_1)^*x_1x_2^*\right)\left(e \cup x_1 \cup x_2(x_2 \cup x_1x_1)^*x_1(e \cup x_2)\right)$. (б) $x_1^*x_2(x_2x_1^*x_2)^*\left(e \cup x_1\left((x_2x_1^*x_2)^*x_1\right)^*(x_2x_1^*x_2)^*\right)$. **2.** Умови задачі задовольняють такі автомати: $A_1 = (B, U, \delta, a_1, F)$, $A_2 = (B, U, \delta, a_1, U \backslash F)$. **3.** Спочатку слід побудувати скінченний автомат $A = (X, U, \delta, a_1, F)$, що зображує подію P. Шуканим автоматом буде $A' = (X, U, \delta, a_1, U \backslash F)$. **4.** (а) $b(a \cup b)^*a$. (б) $(a \cup b)^*aa(a \cup b)^*a$. (в) $bb(a \cup b)^*$. (г) $\varnothing$. (д) a. (е) $\varnothing$. (є) $(ab^*)^*a$. (ж) $(ba \cup bb)(aa \cup ab \cup ba \cup bb)^*$. **5.** (а) $0(00 \cup 01 \cup 10 \cup 11)^*$. (б) Нехай $r = (00 \cup 11)^*$. Тоді $q = r\left((01 \cup 10)r(01 \cup 10)\right)^*r$ задає подію, що складається зі слів у алфавіті $\{0, 1\}$, які містять парну кількість нулів і одиниць. Шуканий регулярний вираз має вигляд $q1q$. (в) $q0q$ (див. (б)). **6.** Множина станів $U = \{1, 2, 3, 4\}$, 1 — початковий, 3 — заключний стани. Функція переходів задається такими співвідношеннями: $\delta(1, a) = \delta(2, a) = \delta(4, a) = \delta(4, b) = 4$, $\delta(1, b) = 2$, $\delta(2, b) = \delta(3, a) = \delta(3, b) = 3$. **7.** Регулярна подія P, зображувана в скінченному автоматі A з n станами, є нескінченною тоді й тільки тоді, коли в графі автомата A є маршрут M, що веде з початкового стану a_1 у деякий заключний стан a_f і містить у собі цикл Z. У свою чергу, маршрут M і цикл Z містять у собі відповідно простий ланцюг L, що веде з a_1 в a_f, і простий цикл C, який має з L спільну вершину. Довжини L і C не перевищують $n - 1$, тому в графі автомата A існуватиме маршрут, що відповідає слову w з P, довжина якого задовольняє нерівності $n \leq |w| < 2n$. Навпаки, нехай дана подія P містить слово w довжини, не меншої n. Тоді у відповідному маршруті M деяка вершина (стан) з'являється більше одного разу, і він матиме вигляд $a_1, (a_1, b_1), b_1, (b_1, b_2), \ldots, b_i, \ldots, b_i, (b_i, b_{i+1}), \ldots, a_f$. Нехай p_1 — слово, що відповідає ділянці маршруту M, яка веде з a_1 в b_i, p — слово, що відповідає переходу з b_i в b_i, і p_2 — слово, яке забезпечує перехід із b_i в a_f. Тоді всі слова $p_1p^kp_2$, $k = 0, 1, 2, \ldots$, переводять автомат A з початкового стану a_1 у заключний стан a_f, тому належатимуть P. Отже, подія P нескінченна. **8.** Слід перевірити всі слова w у вхідному алфавіті заданого автомата A, довжини яких містяться в діапазоні від n до $2n$. Подія, зображувана в A, буде нескінченною в тому й тільки тому разі, коли принаймні одне з цих слів переводитиме автомат A з початкового стану в заключний (див. попередню задачу). **9.** Див. доведення достатності в задачі 7.12.7. **10.** (а) Припустимо, що подія P регулярна. Тоді вона задовольняє "лему про накачку", і для достатньо великого k слово $w = a^kb^k$ можна подати у вигляді $u = p_1pp_2$ так, що $p_1p^np_2 \in P$ для всіх $n = 0, 1, 2, \ldots$. Слово u можна записати так: $a^l(a^{k-l}b^m)b^{k-m}$, де $p_1 = a^l$, $p = a^{k-l}b^m$, $p_2 = b^{k-m}$, $l, m = 0, 1, 2, \ldots, k$. Однак для будь-яких припустимих значень l і m слова $a^l(a^{k-l}b^m)^nb^{k-m}$ для всіх $n = 0, 1, 2, \ldots$ не можуть мати вигляд a^lb^l, тобто належати події P. (б)–(є) Див. (а). (ж) Подія P задовольняє "лему про накачку", коли $p_1 = a^{k-1}$, $p = a$ та $p_2 = b^l$. Припустимо, що P — регулярна

та зображувана в автоматі A з n станами. Розглянемо слово $w = a^k b^l$ цієї події, у якому $k > l > n$. Тоді існуватимуть такі значення t і s $(0 \le t < s \le n)$, що слова a^t і a^s переводять автомат A з початкового стану в той самий стан. Це означатиме, що разом з w події P мають належати слова $u_m = a^{k-(s-t)m} b^l$ для $m = 1, 2, \ldots$ і $k > (s-t)m$. Отже, існуватиме значення m, для якого $u_m \in P$, $u_m = a^r b^l$ і $r \le l$, що суперечитиме означенню події P. **11.** (а) Припустімо, що подія P регулярна і зображувана в скінченному автоматі A. Розглянемо послідовність станів b_p, у які A переходить із початкового стану при подачі на його вхід слів a^{p_i}, де $p_1, p_2, \ldots, p_n, \ldots$ — послідовність усіх простих чисел. Остання послідовність нескінченна, тому існуватимуть такі значення j і l $(j < l)$, що $b_j = b_l$. Тоді всі слова $u_m = a^{p_j + (p_l - p_j)m}$ переводитимуть автомат A з початкового стану в стан b_j, який за означенням події P не є заключним, тому $u_m \notin P$ для $m = 1, 2, \ldots$. Однак для $m = p_j$ слово $u_m \in P$. Отримана суперечність спростовує припущення про регулярність події P. (б) Будь-яке слово події P можна записати як a^{ml}, де m — просте число, а $l = 2, 3, \ldots$. Для $l > 2$ кожне таке слово можна перетворити до вигляду $p_1 p p_2$, де $p_1 = p = p^m$, $p_2 = a^{m(l-2)}$ і $p_1 p^k p_2$ належатиме P для всіх $k = 0, 1, 2, \ldots$. **12.** Неважко довести виконання всіх трьох властивостей еквівалентності для відношення R. Для доведення другого твердження припустімо, що подія P регулярна. Тоді існує автомат $A = (X, U, \delta, a_1, F)$, що зображує P. Розглянемо відношення R' на множині X^*: $(v, w) \in R' \Leftrightarrow \delta(a_1, v) = \delta(a_1, w)$. R' — еквівалентність і існує взаємно однозначна відповідність f між фактор-множиною X^*/R' і множиною станів U автомата A. А саме, класу еквівалентності $[v] \in X^*/R'$ f ставить у відповідність стан $a_i \in U$, у який автомат A переходить із початкового стану a_1 під дією всіх слів із $[v]$. Отже, X^*/R' — скінченна множина. Доведемо, що $R' = R$: $(v, w) \in R' \Leftrightarrow \delta(a_1, v) = \delta(a_1, w) \Leftrightarrow \delta(a_1, vt) = \delta(a_1, wt)$ для будь-якого $t \in X^*$ $\Leftrightarrow$ $\big((\delta(a_1, vt) \in F$ і $\delta(a_1, wt) \in F\big)$ або $\big(\delta(a_1, vt) \notin F$ і $\delta(a_1, wt) \notin F\big)$ для $t \in X^*\big)$ $\Leftrightarrow$ $((vt \in P$ і $wt \in P)$ або $(vt \notin P$ та $wt \notin P)$ для $t \in X^*\big)$ $\Leftrightarrow (v, w) \in R$. Навпаки, припустімо, що для даного відношення R фактор-множина X^*/R — скінченна. Розглянемо скінченний автомат $A = (X, X^*/R, \delta, [e], F)$, у якому функцію переходів означено співвідношенням $\delta([v], x) = [vx]$, а F складають класи еквівалентності, що містять усі слова події P. Оскільки $\delta([e], w) = [w]$, то автомат A допускає всі слова з P й тільки ці слова, тобто зображує подію P. Із теореми 7.12 випливає, що подія P — регулярна. **13.** (а) Якщо R_1 і R_2 — регулярні події в алфавіті X, то $R_1 \backslash R_2 = R_1 \cap \overline{R_2}$. Множина регулярних подій замкнена щодо операцій доповнення і перетину, тому $R_1 \backslash R_2$ — регулярна подія. (б) $R_1 \Delta R_2 = (R_1 \backslash R_2) \cup (R_2 \backslash R_1)$ (див. (а)). **14.** Твердження можна довести індукцією за означенням регулярної події. При цьому слід обгрунтувати і використати такі співвідношення: $\mathrm{pr}_i(P_1 \cup P_2) = \mathrm{pr}_i(P_1) \cup \mathrm{pr}_i(P_2)$, $\mathrm{pr}_i(P_1 P_2) = \mathrm{pr}_i(P_1)\mathrm{pr}_i(P_2)$, $\mathrm{pr}_i(P_1^*) = (\mathrm{pr}_i P_1)^*$,
$i = 1, 2$, P_1 і P_2 — події в алфавіті X. **15.** (а) $R = \{a^m b^n \mid m, n = 0, 1, 2, \ldots\}$ — регулярна подія, задана регулярним виразом $a^* b^*$, $Q = \{ a^k b^l \mid k > l \}$ — нерегулярна подія (див. 7.12.10(ж)). $R \cup Q = R$. (б) $R = \{ab\}$, $Q = \{a^k b^k \mid k = 0, 1, 2, \ldots\}$ (див.7.12.10(а)). **16.** Оскільки скінченна подія S — регулярна (див. теореми 7.7 і 7.12), а множина регулярних подій замкнена відносно операцій об'єднання і різниці, то події $R \cup S$ і $R \backslash S$ — регулярні. **17.** (а) Припустимо, що подія $Q \cup S$ — регулярна і автомат A зображує цю подію. Вилучимо з множини заключних станів автомата A стани, у які ведуть слова зі скінченної множини $S \backslash Q$. Отримаємо автомат A', що зображує подію Q, а цей висновок суперечить умові про нерегулярність Q. (б) Припустимо, що $Q \backslash S$ — регулярна подія. Тоді подія $(Q \backslash S) \cup S = Q \cup S$ також регулярна (див. 7.12.16(а)). Отриманий висновок суперечить твердженню попереднього пункту. **18.** Справедливість твердження випливає із замкненості множини регулярних подій відносно теоретико-множинних операцій об'єднання, перетину і доповнення. Отже, виконуватимуться всі аксіоми булевої алгебри (див. розд. 2.7).

7.15. 4. Якщо A_1 і A_2 — скінченні автомати, що зображують події R_1 і R_2 відповідно, то паралельне сполучення автоматів A_1 і A_2 зображує подію $R_1 \times R_2$.

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. *Александров П. С.* Введение в теорию множеств и общую топологию. — М.: Наука, 1977. — 368 с.
2. *Андерсон Д. А.* Дискретная математика и комбинаторика.– М.: Вильямс, 2003. — 960 с.
3. *Бардачов Ю. М., Соколова Н. А., Ходаков В. Є.* Дискретна математика. — К.: Вища шк., 2002. — 287 с.
4. *Виленкин Н. Я.* Комбинаторика. — М.: Наука, 1969. — 344 с.
5. *Гаврилов Г. П., Сапоженко А. А.* Задачи и упражнения по курсу дискретной математики. — М.: Наука, 1992. — 384 с.
6. *Гаврилов Г. П., Сапоженко А. А.* Сборник задач по дискретной математике. — М.: Наука, 1977. — 368 с.
7. *Глушков В. М.* Введение в кибернетику. — К.: Изд-во АН УССР, 1964. — 324 с.
8. *Глушков В. М.* Синтез цифровых автоматов. — М.: Физматгиз, 1962. — 476 с.
9. *Глушков В. М., Цейтлин Г. Е., Ющенко Е. Л.* Алгебра, языки, программирование. — 3-е изд., перераб. и доп. — К.: Наук. думка, 1989. — 376 с.
10. *Дискретная* математика и математические вопросы кибернетики. — Т.1 / Под общ. ред. С. В. Яблонского, О. Б. Лупанова. — М.: Наука, 1974. — 312 с.
11. *Єжов І. І., Скороход А. В., Ядренко М. Й.* Елементи комбінаторики. — К.: Вища шк., 1972. — 84 с.
12. *Зыков А. А.* Основы теории графов. — М.: Наука, 1987. — 381 с.
13. *Калужнин Л. А.* Введение в общую алгебру. — М.: Наука, 1973. — 448 с.
14. *Калужнін Л. А., Королюк В. С.* Алгоритми і математичні машини. — К.: Вища шк., 1964. — 284 с.
15. *Клини С.* Математическая логика. — М.: Мир, 1973. — 480 с.
16. *Кристофидес Н.* Теория графов. Алгоритмический подход. — М.: Мир, 1978. — 432 с.
17. *Кузнецов О. П., Адельсон-Вельский Г. М.* Дискретная математика для инженера. — 2-е изд., перераб. и доп. — М.: Энергоатомиздат, 1988. — 480 с.
18. *Кук Д., Бейз Д.* Компьютерная математика. — М.: Наука, 1990. — 384 с.
19. *Куратовский К., Мостовский А.* Теория множеств. — М.: Мир, 1970. — 318 с.
20. *Лавров И. А., Максимова Л. Л.* Задачи по теории множеств, математической логике и теории алгоритмов. — М.: Наука, 1975. — 240 с.
21. *Лекции* по теории графов / В. А. Емеличев, О. И. Мельников, В. И. Сарванов, Р. И. Тышкевич — М.: Наука, 1990. — 384 с.
22. *Мендельсон Э.* Введение в математическую логику. — М.: Мир, 1976. — 320 с.
23. *Оре О.* Теория графов. — М.: Наука, 1980. — 336 с.
24. *Рейнгольд Э., Нивергельт Ю., Део Н.* Комбинаторные алгоритмы. — М.: Мир, 1980. — 476 с.
25. *Свами М., Тхуласираман К.* Графы, сети, алгоритмы. — М.: Мир, 1984. — 454 с.
26. *Столл Р.* Множества. Логика. Аксиоматические теории. — М.: Просвещение. — 1968. — 231 с.
27. *Трохимчук Р. М.* Збірник задач з дискретної математики (розділ “Множини і відношення”). — К.: РВЦ КУ, 1997. — 69 с.
28. *Трохимчук Р. М.* Збірник задач з теорії булевих функцій: Навч. посіб. — К.: ВПЦ КУ, 2002. — 104 с.
29. *Трохимчук Р. М.* Збірник задач із теорії графів: Навч. посіб. — К.: РВЦ КУ, 1998. — 57 с.
30. *Трохимчук Р. М.* Основи дискретної математики: Практикум. — К.: МАУП, 2004. — 136 с.
31. *Уилсон Р.* Введение в теорию графов. — М.: Мир, 1977. — 207 с.
32. *Френкель А., Бар-Хиллел И.* Основания теории множеств. — М.: Мир,1966. — 556 с.
33. *Харари Ф.* Теория графов. — М.: Мир, 1973. — 300 с.

34. *Хромой Я. В.* Збірник задач і вправ з математичної логіки. — К.: Вища шк., 1978. — 160 с.
35. *Хромой Я. В.* Математична логіка. — К.: Вища шк., 1983. — 208 с.
36. *Шиханович Ю. А.* Введение в современную математику. — М.: Наука. — 376 с.
37. *Шрейдер Ю. А.* Равенство, сходство, порядок. — М.: Наука. — 1971. — 254 с.
38. *Яблонский С. В.* Введение в дискретную математику. — М.: Наука, 1979. — 272 с.

ПРЕДМЕТНИЙ ПОКАЖЧИК

ЗМІСТ

The bases of discrete (computer) mathematics as a component of theoretical (computer) cybernetics — sciences that is engaged in development and researches of computer and program system was considered at the educational manual. The material was illustrated by many examples. Tasks and plans were presented to each part that is able to use this manual for practical studies and self-work. To the majority tasks were produced answers or explanations.

The manual is intended for high educational institutions and everyone, who is eager to receive fundamental knowledge in the area of modern information science.

Навчальне видання
Трохимчук Ростислав Миколайович
ДИСКРЕТНА МАТЕМАТИКА
Навчальний посібник

Educational edition
Trokhymchuk, Rostyslav M.
DISCRETE MATHEMATICS
Educational manual

Відповідальний редактор *С. Г. Рогузько*
Редактор *С. Г. Атдаєва*
Коректори: *А. А. Карпова, Т. М. Федосенко*
Комп'ютерне верстання *І. І. Савіцький, Т. Г. Замура*
Оформлення обкладинки *С. В. Фадєєв*

Підп. до друку 22.12.09. Формат 60x84 $^1/_{16}$. Папір офсетний. Друк офсетний.
Ум. друк. арк. 30,69. Обл-вид. арк. 30,8. Тираж 1300 пр. Зам. № 56.

Міжрегіональна Академія управління персоналом (МАУП)
03039 Київ-39, вул. Фрометівська, 2, МАУП

ДП «Видавничий дім «Персонал»
03039 Київ-39, просп. Червонозоряний, 119, літ. XX

Свідоцтво про внесення до Державного реєстру суб'єктів видавничої справи
ДК № 8 від 23.02.2000

Дискретна математика

ДП «Видавничий дім
«ПЕРСОНАЛ»
м. Київ

Виготовлення та реалізація
книжкової продукції

Виготовлення рекламної
та поліграфічної продукції

Видавництво: (044)525-61-85

Відділ реалізації: 490-95-17
E-mail: opt@iapm.edu.ua

Відділ реклами: 490-95-18
E-mail: personalmarket@iapm.edu.ua

Факс: 490-95-19

ISBN 978-966-608-863-8