МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

ДИСКРЕТНА МАТЕМАТИКА

КОНСПЕКТ ЛЕКЦІЙ

Частина 2

Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського як навчальний посібник для студентів, які навчаються за спеціальністю 113 «Прикладна математика», освітньою програмою «Наука про дані та математичне моделювання»

Київ

КПІ ім. Ігоря Сікорського

2019

Рецензенти: *Романкевич В.О.*, д-р техн. наук., доц., професор кафедри системного програмування і спеціалізованих комп'ютерних систем ФПМ

Жабіна В.В., канд. техн. наук., доц., доцент кафедри програмного забезпечення комп'ютерних систем ФПМ

Відповідальний редактор *Сулема Є.С.*, канд. техн. наук, доц.

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 3 від 28.11.2019 р.)
за поданням Вченої ради факультету прикладної математики (протокол № 5 від 11.11.2019 р.)

Електронне мережне навчальне видання

Темнікова Олена Леонідівна

ДИСКРЕТНА МАТЕМАТИКА
КОНСПЕКТ ЛЕКЦІЙ
ЧАСТИНА 2

Дискретна математика: Конспект лекцій (Частина 2) [Електронний ресурс]: навч. посіб. для студ. спеціальності 113 «Прикладна математика», освітньої програми «Наука про дані та математичне моделювання» / О.Л.Темнікова ; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 1,84 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2019. – 128 с.

Навчальний посібник розроблено для ознайомлення студентів з лекційним курсом другого кредитного модулю з дисципліни «Дискретна математика» та призначено для студентів, які навчаються за спеціальністю 113 «Прикладна математика», освітньою програмою «Наука про дані та математичне моделювання» факультету прикладної математики КПІ ім. Ігоря Сікорського. «Дискретна математика: Конспект лекцій (Частина 2)» містить викладання основ теорії булевої алгебри: побудова таблиць істинності, алгебраїчні перетворення, властивості булевих функцій, повнота систем, різні методи мінімізації. Розглядаються теорія висловлювань та формальні системи, як моделі булевої алгебри. Також надаються основи теорії абстрактних автоматів та мереж Петрі.

ЗМІСТ

ВСТУП

Навчальне видання призначене для студентів, які вивчають другий кредитний модуль навчальної дисципліни «Дискретна математика», який є однією з частин дисципліни «Дискретна математика» з циклу загальної підготовки навчального плану спеціальності 113 «Прикладна математика» і є базовим при навчанні за освітньою програмою «Наука про дані та математичне моделювання», що спрямована на розроблення і використання нових інформаційних технологій, моделювання складних систем, систем штучного інтелекту, і для вирішення інших задач професійної діяльності. Лекційний курс з кредитного модулю «Дискретна математика (Частина 2)» розраховано на 36 академічних годин аудиторних занять, вивчається на факультеті прикладної математики у 2 семестрі.

Метою викладання кредитного модулю є оволодіння основними поняттями і методами, що дає змогу сформувати у студентів компетенції, потрібні для вивчення наступних дисциплін спеціальності, формування світогляду на дискретну математику як на фундаментальну науку, що призначена для формалізації знань, розробки і використання нових інформаційних технологій, моделювання складних систем, систем штучного інтелекту.

Після засвоєння кредитного модуля студенти мають знання: основних понять булевої алгебри; принципів дії абстрактих автоматів та мереж Петрі; математичних методів дискретного аналізу в прикладних науках і розв'язанні практичних задач; та уміння: досліджувати системи булевих функцій на повноту, замкнутість, несуперечність, знаходити досконалі і мінімальні стандартні форми; синтезувати та аналізувати моделі, побудовані на базі абстрактних автоматів та мереж Петрі.

Навчальне видання має стати в нагоді під час опрацювання матеріалів лекцій, підготовки до екзамену, до практичних занять.

1. БУЛЕВА АЛГЕБРА

Засновником формальної логіки вважається Арістотель, хоча основну увагу він приділив розгляду силогізмів. Протягом двох тисячоліть логіки в основному займалися вивченням силогізмів, і ще в 1797 р. Іммануїл Кант писав про логіку як про «замкнуту і завершену теорію».

Поворотний пункт в історії логіки настав в 1847р., коли Джордж Буль (1815-1864), вчений-самоучка, син бідного шевця, опублікував статтю «Математичний аналіз логіки». Ця та деякі інші статті принесли Булю популярність, і він, не маючи наукового ступеня, отримав пропозицію зайняти посаду професора математики в Куїнз-Коледж ірландського міста Корка. Там він написав свій знаменитий трактат «Дослідження законів мислення», що вийшов друком у 1854р. в Лондоні. Основна ідея праці - заміна всіх слів, що вживаються в формальній логіці, символами - висловлювалася і до Буля, але він був першим, кому вдалося розробити практично придатну систему. Недоліки позначень Буля були обумовлені прагненням їх автора зробити систему якомога більше схожою на традиційну алгебру. В даний час, кажучи про булеву алгебру, математик має на увазі абстрактну систему символів, властивості якої описуються аксіоматично.

1.1. Абстрактна булева алгебра

Згадаймо, що алгеброю називається будь-яка множина об'єктів А (чисел, функцій, матриць, геометричних об'єктів та ін.), на якій задана деяка система кінцево-місцевих операцій, тобто функцій, змінні яких приймають значення з А, і самі функції приймають значення в А.

Булеву алгебру, так само як і всі інші абстрактні алгебри можна інтерпретувати безліччю різних способів. Сам Буль інтерпретував свою систему в дусі Арістотеля як алгебру класів і їх відносин. Оскільки математики давно відмовилися від початкових позначень, прийнятих в роботі Буля, булеву алгебру в даний час прийнято записувати в позначеннях теорії множин.

Визначення булевої алгебри було дано у темі «Решітки» в першому семестрі:

Булевою алгеброю $B = \langle L, \vee, \wedge, ', \mathbf{0}, \boldsymbol{I} \rangle$ називається алгебра з двома бінарними операціями $\vee$ і $\wedge$, однією унарною операцією $'$ і двома нулярними операціями $\mathbf{0}$ і $\boldsymbol{I}$, що задовольняють твердженням **L1 — L9**. *Нулярні* операції *виділяють* елементи $\mathbf{0}, \boldsymbol{I}$ множини L, ці элементи називаються *виділеними елементами*.

$x \wedge x = x, \quad x \vee x = x$	(*ідемпотентність*)	**L1**
$x \vee y = y \vee x, \quad x \wedge y = y \wedge x$	(*комутативність*)	**L2**
$x \wedge (y \wedge z) = (x \wedge y) \wedge z,$ $x \vee (y \vee z) = (x \vee y) \vee z$	(*асоціативність*)	**L3**
$x \wedge (x \vee y) = x, \quad x \vee (x \wedge y) = x$	(*поглинання*)	**L4**
якщо $x \le z$, то $x \vee (y \wedge z) = (x \vee y) \wedge z$ якщо $z \le x$, то $x \wedge (y \vee z) = (x \wedge y) \vee z$	*(модулярність)*	**L5**
$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$	*(дистрибутивність)*	**L6**
$x \wedge x' = \mathbf{0}, \qquad x \vee x' = \boldsymbol{I};$	*(протиріччя та виключення третього)*	**L7**
$(x')' = x$	(*інволюція*)	**L8**
$(x \wedge y)' = x' \vee y', \qquad (x \vee y)' = x' \wedge y'$	(*закони де Моргана*)	**L9**

Оскільки будь-яку решітку можна розглядати як алгебру з двома операціями, булеву решітку, в якій кожен елемент має єдине доповнення, можна розглядати як булеву алгебру з трьома операціями. Ми розглянемо в деякому сенсі «мінімальну» булеву алгебру, яка будується на решітці, що складається з двох елементів. Двійкові змінні можуть мати різні інтерпретації, наприклад, положення вимикача «розімкнуто»-«замкнуто» в теорії релейно-контактних схем, «так»-«ні», або «істинно»-«хибно». Остання інтерпретація використовується в алгебрі висловлювань і алгебрі предикатів, в алгоритмічних мовах (Паскаль, Пролог), де логічні змінні можуть приймати одне з можливих значень: true (істинно) або false (хибно). У булевій алгебрі прийнято використовувати ***0*** - true (істинно) або ***1*** - false (хибно).

У абстрактної булевої алгебри природа множини {0, 1} не розглядається. Решіткові операції об'єднання ($\vee$), перетину ($\wedge$) і доповнення ($\neg$) мають інші найменування і розглядаються як операції алгебри. Всі властивості булевих решіток, зрозуміло, виконуються в булевій алгебрі і визначаються в ній як аксіоми і теореми.

Взагалі, будь-яка множина елементів із заданими на ньому двома двомісними операціями, однією одномісною і виконуваними аксіомами комутативності, асоціативності, дистрибутивності, ідемпотентності, поглинання, інволюції, законів де Моргана називається *булевою алгеброю.*

Булева алгебра називається виродженою, якщо ***I*** та ***0*** співпдають. В такому випадку з огляду на рівності $\boldsymbol{x = x \wedge I = x \wedge 0 = 0}$ вона не містить ніяких інших елементів, а значить складається рівно з одного елемента. Будь-яка невироджена булева алгебра (а тільки такі ми і будемо розглядати) містить два нейтральних елементи - нульовий ***0*** і одиничний ***I***.

Характерною рисою комп'ютерів є зведення всіх обчислювальних структур (чисел, символів, масивів і т.п.) до двійкових слів і алгоритмів обробки. З математичної точки зору ми будемо мати справу з окремим випадком булевих алгебр.

Визначення 1.1. Будь-яку змінну, яка може приймати одне з двох можливих значень 0 або 1, назвемо булевою змінною.

Визначення 1.2. Функції, які визначені на множині {0, 1} і приймають значення з цієї множини, називаються булевими функціями.

Булева функція називається n-місною, якщо число її змінних дорівнює n. Занумеруємо змінні булевої функції. Сукупність значень змінних булевої функції назвемо набором (або кортежем). Набір довжини n будемо називати n-набором. Тоді областю визначення n-місної булевої функції є сукупність всіляких упорядкованих n-наборів, компонентами яких є значення булевих змінних 0 або 1.

Кожен n-набір можна розглядати як запис деякого числа в двійковій системі числення. Нехай n-набір має вигляд: $a_1, a_2, \dots, a_n$, де a_i=0 або 1.Тоді число, визначене як $a_1*2^{n-1}+a_2*2^{n-2}+\dots+a_n$, назвемо номером даного набора. Зрозуміло, що номером n-набора може бути будь-яке натуральне число $0<=K<=2^{n-1}$.

Наприклад, для булевої функції $f(x, y, z)$ множина значень $x=1$, $y=0$, $z=1$ записується як набор 101. Булева функція може бути задана за допомогою таблиці – перерахування її значень на всіх наборах. Множину наборів прийнято записувати в лексикографічному порядку, тому кожен набір являє собою код двійкового числа. Відповідне йому десяткове число будемо називати номером набору. Наприклад, номер набору 101 дорівнює 5, номер набору 110 - 6.

Твердження 1.1. Кількість наборів булевої функції $f(x_1, \dots, x_n)$ від n змінних дорівнює 2^n. Кількість булевих функцій від n змінних дорівнює 2^{2^n}.

Доведення цього твердження безпосередньо випливає з основних співвідношень кардинальної арифметики. Дійсно, множина всіх наборів булевої функції від n змінних утворена декартовим добутком $\{0, 1\}^n$, потужність якого дорівнює 2^n. Множина всіх булевих функцій від n змінних є множина відображень $\{0, 1\}^n \to \{0, 1)$, потужність якої дорівнює 2^{2^n}.

Доведення також легко проводитися методом математичної індукції (по n). При n = 1 теорема очевидна. Припустимо, вона справедлива при n = k, тобто 2^k для k-наборів. Візьмемо довільно один з них. Приписуючи до нього справа 0 або 1, отримуємо (k + 1) - набір. Очевидно, що таким чином кожен (k + 1) - набір виходить з деякого k-набору і, що, якщо два k-набору різні, то і всі отримані з них (k + 1) - набори теж різні. Т.ч. існує $(2^k)*2=2^{k+1}$ для (k+1)-набору, що й треба було довести.

Булева функція може бути визначена не на всіх наборах. В цьому випадку вона називається частково визначеною. Будь-яку частково визначену булеву функцію можна довизначити (взагалі кажучи, довільним чином) на тих наборах, на яких вона раніше не була визначена. Надалі будемо говорити про усюди визначених булевих функціях.

Існує 4 булеві функції від однієї змінної та 16 булевих функцій від двох змінних, тримісних - вже 256, а 4-х-місних - близько 65 тисяч.

Всі набори розташуємо в порядку зростання їх номерів, а поруч з кожним набором розмістимо значення булевої функції на цьому наборі. Отриману таким чином таблицю назвемо таблицею значень, або таблицею істинності булевої функції. У таблиці 1.1 наведено чотири булеві функції від однієї змінної:

f_0 – константа 0,

f_1 – тотожна функція $f(x) = x$,

f_2 – заперечення $f(x) = \neg x$,

f_3 – константа 1.

Таблиця 1.1.

x	f_0	f_1	f_2	f_3
0	0	0	1	1
1	0	1	0	1

В табл. 1.2 приведені функції від двох змінних.

Таблица.1.2.

x	y	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{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

Ці функції:

f_0 – константа 0

f_1 – $x \wedge y$ (або x & y) – кон'юнкція

f_2 – Заперечення імплікації $x \to y$ $f(x,y) = \neg (x \to y)$

f_3 – Тотожна функція $f(x,y) = x$

f_4 – Заперечення імплікації $y \to x$ $f(x,y) = \neg (y \to x)$

f_5 – Тотожна функція $f(x,y) = y$

f_6 – $x \oplus y$ – сума за *mod* 2 (операція Жегалкіна)

f_7 – $x \vee y$ – диз'юнкція[1]

f_8 – заперечення диз'юнкції $f(x,y) = \neg (x \vee y) = (x \downarrow y)$ – стрілка Пірса (іноді можна зустріти назву функція Даггера або функція Вебба)

f_9 – $x \equiv y$ (або $x \sim y$) – еквівалентність (еквіваленція)

f_{10} – заперечення $f(y) = \neg y$

f_{11} – $y \to x$ (або $y \supset x$) – імплікація

f_{12} – заперечення $f(x) = \neg x$

f_{13} – $x \to y$ (або $x \supset y$) – імплікація

f_{14} – заперечення кон'юнкції $f(x,y) = \neg (x \wedge y) = (x \mid y)$ – штрих Шеффера

f_{15} – константа 1

Розглядаючи таблицю булевих функцій від двох змінних, можна помітити, що деякі з функцій не залежать від однієї з змінних, наприклад, $f_3(x,y)=\neg x$ не залежить від змінної y, функція $f_{10}(x,y)=\neg y$ не залежить від x, $f_{15}(x,y)=1$ не залежить від x і y.

[1] *Об'єднання класів Буль записував як x + y і розумів в «винятковому» сенсі, що відповідає сучасній операції x⊕y - додавання за mod 2 (операція Жегалкіна). Об'єднання «у включному» сенсі або диз'юнкцію ввів англійський логік і економіст Джевонс. Інтерпретація Джевонса виявилася зручнішою та стала згодом загальноприйнятої.*

Визначення. Змінна x_i функції $F(x_1,\ldots,x_i,\ldots,x_n)$ називається несуттєвою (фіктивною), якщо x_i в будь-якому наборі змінних $x_1,\ldots,x_{i-1},x_i,\ldots, x_n$ не змінює значення функції F, тобто, якщо виконується рівність $F(x_1,\ldots, x_{i-1},x_i,x_{i+1},\ldots, x_n)= F(x_1,\ldots, x_{i-1},\neg x_i,x_{i+1},\ldots, x_n)$ для будь-яких значень змінних $x_1,\ldots, x_{i-1},x_{i+1},\ldots, x_n$. В іншому випадку змінна називається суттєвою.

Якщо деяка булева функція $f(x_1,\ldots,x_i,\ldots,x_n)$ містить фіктивну змінну x_i, то по суті вона залежить від (n-1)-ї змінної. Функція $g(x_1,\ldots, x_{i-1},x_{i+1},\ldots, x_n)$ отримана з функції f шляхом видалення фіктивної змінної, еквівалентна вихідній функції f.

Введення k фіктивних змінних в булеву функцію $f(x_1,\ldots,x_n)$ робить її формально залежною від (n + k) змінних і дає функцію $g(x_1,\ldots,x_n,x_{n+1},\ldots,x_{n+k})$, еквівалентну вихідній.

Твердження 1.2. Множина, що складається з двох значень 0 і 1, на якому визначені унарна операція заперечення $\neg$ згідно табл. 1.1 і бінарні операції диз'юнкції $\vee$ і кон'юнкції $\wedge$ відповідно до табл. 1.2, є булевою алгеброю.

Істинність твердження 1.2 випливає з теорії решіток: операція заперечення булевої алгебри визначається як доповнення в решітці-2, диз'юнкції $\vee$ і кон'юнкції $\wedge$ – як об'єднання і перетин відповідно.

Операція заперечення при рукописному написанні зазвичай записується як штрих над символом змінної, наприклад: $\bar{x}, \overline{(x \vee y)}$, проте, ми будемо використовувати символ $\neg$, наприклад: $\neg x$, $\neg(x \vee y)$. Операція кон'юнкції в булевій алгебрі зазвичай позначається символом $\wedge$, в алгебрі висловлювань – символом &, а $x \wedge y$ записується просто як xy.

1.2 Основні аксіоми й теореми булевої алгебри

Аксіоми і теореми булевої алгебри безпосередньо випливають з властивостей булевих решіток. Наведемо список аксіом і основних теорем булевої алгебри.

Аксіоми булевої алгебри:

(1.1) *комутативні закони:*

$a \vee b = b \vee a$;

$a \wedge b = b \wedge a$;

(1.2) *дистрибутивні закони:*

$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$;

$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$;

(1.3) *асоціативні закони:*

$\mathrm{a} \vee (b \vee c) = (a \vee b) \vee c$;

$a \wedge (b \wedge c) = (a \wedge b) \wedge c$;

(1.4) *властивості 0 й 1:*

$a \vee 0 = a$, $a \wedge 1 = a$;

(1.5) *закон виключення третього:*

$a \vee \neg a = 1$; та *закон протиріччя:* $a \wedge \neg a = 0$.

Теореми:

(1.6) якщо для всіх a $a \vee b = a$, то $b = 0$;

якщо для всіх a $a \wedge b = a$, то $b = 1$;

(1.7) якщо $a \vee b = 1$ і $a \wedge b = 0$, то $b = \neg a$;

(1.8) $\neg\neg a = a$;

(1.9) $\neg 0 = 1$, $\neg 1 = 0$;

(1.10) *закон ідемпотентності:*

$a \vee a = a$;

$a \wedge a = a$;

(1.11) $a \vee 1 = 1$, $a \wedge 0 = 0$;

(1.12) *закони поглинання:*

$a \vee (a \wedge b) = a$;

$a \wedge (a \vee b) = a$;

(1.13) *закони де Моргана:*

$\neg(a \vee b) = \neg a \wedge \neg b$;

$\neg(a \wedge b) = \neg a \vee \neg b$;

(1.14) *закони склеювання:*

$(a \vee b) \wedge (a \vee \neg b) = a$;

$(a \wedge b) \vee (a \wedge \neg b) = a$;

(1.15) *закони поглинання доповнення:*

$a \vee (\neg a \wedge b) = a \vee b$;

$a \wedge (\neg a \vee b) = a \wedge b$.

Це найбільш проста модель булевої алгебри, разом з тим найбільш важлива для комп'ютерних наук. Інша поширена модель - модель обчислення висловлювань.

Якщо взяти деяку непорожню множину **A**, то її множина-ступінь **ρ(A)** буде моделлю булевої алгебри, якщо домовитися про наступне (встановити ізоморфізм між поняттям та визначенням):

- елементи цієї булевої алгебри – різні підмножини множини A;
- операція кон'юнкції визначається як перетин множин, диз'юнкції – об'єднання множин, операція заперечення – абсолютне доповнення;
- множина ʊ (універсум) і ø – відповідно одиниця й нуль алгебри.

Неважко переконатися, що всі аксіоми, наведені вище, будуть виконуватися.

1.3 Булеві формули

Визначення 1.3.

- Кожна змінна є формулою.
- Якщо *x, y* – формули, то формулами є $(\neg x)$, $(x \vee y)$, $(x \,\&\, y)$.
- Інших формул нема.

У зображенні формул прийняті наступні допущення: зовнішні дужки опускають; встановлюють пріоритети виконання операцій в наступному порядку:

$\neg$ – заперечення (найвищій пріоритет),

$\wedge$ – кон’юнкція**,**

$\vee$ – диз'юнкція,

$\rightarrow$, $\equiv$ – імплікація й еквіваленція (мають однаковий пріоритет, тим паче, що вони не є алгебраїчними операціями, а слугують для скорочення запису).

З урахуванням цих пріоритетів надлишкові дужки також опускаються.

Для кожної булевої формули можна побудувати її таблицю істинності, обчислюючи її значення на кожному наборі змінних. Звідси випливає, що кожній булевій формулі відповідає деяка булева функція, а формулу булевої алгебри можна розглядати як представлення булевої функції від деяких змінних.

Кожна булева формула єдиним чином визначає булеву функцію. Єдиність випливає з того, що для кожної формули булевої алгебри можна побудувати єдину таблицю істинності. Тому в подальшому поняття «булева формула» і «булева функція» будемо вживати як рівнозначні.

Протилежне твердження про відповідність булевих формул і функцій, не справедливо, тобто кожній булевій функції можна поставити у відповідність безліч булевих формул.

Булеві формули називаються еквівалентними, або рівносильними, якщо їх таблиці збігаються. Якщо дві рівносильні формули з'єднати за допомогою операції еквівалентності, то отримана формула буде тотожно дорівнювати одиниці. Такі тотожно істинні формули називаються тавтологіями або загальнозначущими. Формули, тотожно рівні нулю, тобто тотожно хибні, називаються суперечливими або суперечностями. Аксіоми і теореми булевої алгебри задають рівносильні формули.

Розглянемо деякі рівносильні формули (див. Табл. 1.3).

Таблиця 1.3.

x	y	$x\to y$	$\neg x\vee y$	$y\to x$	$(x\to y)(y\to x)$	$x\equiv y$	$x\oplus y$	$\neg(x\oplus y)$
0	0	1	1	1	1	1	0	1
0	1	1	1	0	0	0	1	0
1	0	0	0	1	0	0	1	0
1	1	1	1	1	1	1	0	1

Ми бачимо, що таблиці значення формул $x \to y$ й $\neg x \vee y$ збігаються. Отже, ці формули еквівалентні:

(1.16) $x \to y = \neg x \vee y = \neg(x\&\neg y)$.

Еквівалентні також наступні формули:

(1.17) $x \equiv y = \neg(x \oplus y)$,

(1.18) $x \equiv y = (x \to y)(y \to x)$.

В (1.18) ми можемо замінити імплікацію на рівносильну їй формулу (1.16). В результаті отримаємо:

(1.19) $x \equiv y = (x \to y)(y \to x) = (\neg x \vee y)(\neg y \vee x)$ або $(\neg x\&\neg y)\vee(x\&y)$.

(1.20) Тоді, $x \oplus y=\neg(x \equiv y)=(\neg x\&y)\vee(x\&\neg y)=(x\vee y)\&(\neg x\vee\neg y)$.

Ці равносильності пояснюють, чому при визначенні булевої формули використовуються тільки операції $\neg$, $\vee$, $\wedge$ – операції імплікації, еквіваленції й складання по модулю 2 можуть бути введені за допомогою рівносильних формул (1.16) - (1.20).

За допомогою рівносильних формул можна проводити еквівалентні перетворення булевих формул. Наприклад, перетворимо булеву формулу $xy\to(\neg y\oplus z)$.

$\mathbf{x\,y \to (\neg y \oplus z) =}$	$x \to y = \neg x \vee y$
$\mathbf{= \neg(x\,y) \vee (\neg y\,\neg z \vee y\,z) =}$	$x \oplus y=(\neg xy)\vee(x\neg y)$, $\neg\neg y = y$
$\mathbf{= \neg x \vee \neg y \vee \neg y\,\neg z \vee y\,z =}$	*закон де Моргана*
$\mathbf{= \neg x \vee \neg y \vee y\,z = \neg x \vee \neg y \vee z}$	*закон поглинання*
	закон поглинання доповнення

Побудуємо також для вихідної формули і перетвореної таблицю істинності (див. Табл.1.4) і переконаємося в їх равносильності.

Таблиця 1.4.

x	y	z	**xy**	$\neg y$	$\neg y \oplus z$	$xy \to (\neg y \oplus z)$	$\neg x$	$\neg x \vee \neg y$	$\neg x \vee \neg y \vee z$
0	***0***	***0***	0	1	1	1	1	1	1
0	***0***	***1***	0	1	0	1	1	1	1
0	***1***	***0***	0	0	0	1	1	1	1
0	***1***	***1***	0	0	1	1	1	1	1
1	***0***	***0***	0	1	1	1	0	1	1
1	***0***	***1***	0	1	0	1	0	1	1
1	***1***	***0***	1	0	0	0	0	0	0
1	***1***	***1***	1	0	1	1	0	0	1

Важливо вміти визначати, чи є формула тавтологією. Для цього можна скористатися одним з методів: побудувати таблицю істинності (формула повинна тотожно дорівнювати 1), виконати еквівалентні перетворення (і отримати 1) або застосувати метод редукції (зведення до протиріччя – передбачається, що існує набір, на якому формула приймає значення 0, далі намагаються підібрати відповідні значення змінних або спростувати тезу).

1.4. Диз'юнктивні та кон'юнктивні нормальні форми

Оскільки одна і та сама булева функція може бути представлена різноманітними формулами, виникає питання про єдиність представлення булевих функцій, тобто знаходження такої універсальної форми, щоб кожна булева функція була представлена в ній. Такими формами в булевій алгебрі є *досконалі диз'юнктивна та кон'юнктивна нормальні форми*.

Визначення 1.4. Будь-яка булева формула, побудована за допомогою однієї тільки операції диз'юнкції над булевими змінними чи їхніми запереченнями, називається *елементарною диз'юнкцією* чи *диз'юнктом*. Наприклад, $x_1 \vee x_2 \vee \neg x_3 \vee x_4$ – елементарна диз'юнкція.

Визначення 1.5. Будь-яка булева формула, побудована за допомогою однієї тільки операції кон'юнкції над булевими змінними чи їхніми запереченнями, називається *елементарною кон'юнкцією,* чи *кон'юнктом*. Наприклад, $x_1 x_2 \neg x_3 \neg x_4$ – елементарна кон'юнкція.

Теорема 1.1. Для того, щоб елементарна диз'юнкція тотожно дорівнювала одиниці, необхідно та достатньо, щоб до неї входила деяка змінна разом із її запереченням.

Теорема 1.2. Для того, щоб елементарна кон'юнкція тотожно дорівнювала нулю, необхідно та достатньо, щоб у неї входила деяка змінна разом із її запереченням.

Справедливість теорем 1.1, 1.2 випливає з аксіоми (1.4) і визначення операцій диз'юнкції та кон'юнкції.

Визначення 1.6. Булева функція називається *конституентою одиниці* (мінтермом), якщо вона дорівнює одиниці тільки на одному наборі своїх аргументів.

Визначення 1.7. Булева функція називається *конституентою нуля* (макстермом), якщо вона дорівнює нулю тільки на одному наборі своїх аргументів.

Приклад. Серед булевих функцій двох аргументів кон'юнкція та стрілка Пірса є конституентами одиниці, а диз'юнкція, імплікація, штрих Шеффера є конституентами нуля.

Теорема 1.3. Не тотожно хибна елементарна кон'юнкція від n-змінних є конституентою одиниці від n-змінних.

Доведення. Введемо позначення: $x^0 = \neg x$, $x^1 = x$. Позначимо параметр 0 чи 1 через α. Тоді $x^\alpha = 1$, якщо $x = \alpha$, і $x^\alpha = 0$, якщо $x \neq \alpha$. Окрім того, $1^\alpha = \alpha$, $0^\alpha = \neg\alpha$.

Розглянемо елементарну кон'юнкцію виду $x_1^{\alpha 1}x_2^{\alpha 2}\ldots x_n^{\alpha n}$, про яку відомо, що вона не дорівнює нулю тотожно. Із теореми 1.2 випливає, що всі змінні, що входять до неї, різні і їм можна надавати незалежні значення.

Надамо змінній x_1 значення α_1, $x_2 = \alpha_2$ і т.д. до $x_n = \alpha_n$. Тоді елементарна кон'юнкція $x_1^{\alpha 1}x_2^{\alpha 2}x_n^{\alpha n}$ дорівнює одиниці по самому вибору значень x_1, x_2, ..., x_n. Вибір цих значень такий, що якщо змінна x_ι входить в елементарну кон'юнкцію із запереченням, то $\alpha_\iota = 0$, і змінній x_ι також надаємо значення, рівне нулю. Тоді, за визначенням, значення $x^{\alpha i} = \neg x = \neg 0 = 1$. Якщо ж змінна x_ι входить в елементарну кон'юнкцію без заперечення, то $\alpha_\iota = 1$ і $x^1 = x = 1$. Таким чином, буде знайдений набір, на якому дана кон'юнкція буде дорівнювати одиниці. Єдиність такого набора випливає із визначення операції кон'юнкції.

Із теореми 1.3 випливає, що будь-яку конституенту одиниці можна представити у вигляді елементарної кон'юнкції. Для цього необхідно утворити кон'юнкцію всіх її аргументів та розставити заперечення над тими змінними, які рівні нулю на наборі, що перетворює функцію в одиницю.

Аналогічну теорему можна довести для елементарної диз'юнкції.

Теорема 1.4. Не тотожно істинна елементарна диз'юнкція від n-змінних є конституентою нуля від n-змінних.

Доведення надається читачеві.

Із цієї теореми випливає, що будь-яку конституенту нуля можна представити у вигляді елементарної диз'юнкції. Для цього необхідно утворити диз'юнкцію всіх змінних та розставити заперечення над тими змінними, які рівні одиниці на наборі, що перетворює функцію в нуль.

Приклад. Нехай функція $f(x, y, z, v)$ дорівнює одиниці на єдиному наборі 0110. Тоді вона є конституентою одиниці і її можна записати у вигляді: $f(x, y, z, v) = \neg xyz\neg v$.

Визначення 1.8. Диз'юнкція елементарних кон'юнкцій, що не містить двох однакових кон'юнкцій, називається *диз'юнктивною нормальною формою (ДНФ).*

Визначення 1.9. Диз’юнктивна нормальна форма, всі кон’юнкції якої є конституенти одиниці, називається *досконалою диз'юнктивною нормальною формою (ДДНФ).*

Визначення 1.10. Кон'юнкція елементарних диз’юнкцій, що не містить двох однакових диз’юнкцій, називається *кон’юнктивною нормальною формою (КНФ).*

Визначення 1.11. Кон’юнктивна нормальна форма, всі диз’юнкції якої є конституенти нуля, називається *досконалою кон'юнктивною нормальною формою (ДКНФ).*

Теорема 1.5. Будь-яку булеву функцію можна представити у вигляді ДДНФ і ДКНФ.

Доведення. Розглянемо довільну булеву функцію $f(x_1, x_2, ..., x_n)$, що не є константою. Нехай $\alpha_1, \alpha_2, ..., \alpha_k$ – набори, на яких функція приймає значення, рівне одиниці, а $\alpha_1, \alpha_2, ..., \alpha_l$ – набори, на яких функція дорівнює нулю (l=2^n-k). Побудуємо елементарні кон'юнкції $K\alpha_1(x_1, ..., x_n)$, $K\alpha_2(x_1, ..., x_n)$, …, $K\alpha_k(x_1, ..., x_n)$, кожна з яких містить всі n змінних і не є тотожно хибною. Побудуємо ці елементарні кон'юнкції таким чином, щоб $K\alpha_1$ приймала значення, рівне одиниці, на наборі α_1, $K\alpha_2$ – на наборі α_2 і т. д. Згідно теоремі 1.3, кожна з цих кон’юнкцій є конституентою одиниці і на решті наборів приймає значення, рівне нулю. По своїй побудові сукупність всіх цих конституент одиниці описує всі одиниці даної функції. Тому їхня диз’юнкція рівносильна даній функції: $f(x_1,..., x_n) = K\alpha_1(x_1, ..., x_n) \vee K\alpha_2(x_1, ..., x_n) \vee ... \vee K\alpha_k(x_1, ..., x_n)$. За визначенням 1.9 отримана формула є досконалою диз'юнктивною нормальною формою даної функції.

Аналогічно можна показати, що формула $A = K\alpha_1 \,\&\, K\alpha_2 \,\&\, … \,\&\, K\alpha_l$, що є кон’юнкцією всіх конституент нуля на наборах, де функція дорівнює нулю, рівносильна даній функції, і є ДКНФ функції.

Приклад. Нехай булева функція $f(x, y, z)$ задана таблицею 1.5. Побудуємо ДДНФ і ДКНФ даної функції.

Таблиця 1.5.

x	0	0	0	0	1	1	1	1
y	0	0	1	1	0	0	1	1
z	0	1	0	1	0	1	0	1
$f(x, y, z)$	1	0	1	0	1	1	0	0

Для побудови ДДНФ булевої функції необхідно розглянути всі набори, на яких функція дорівнює одиниці, і утворити конституенти одиниці, що відповідають цим наборам. Отримані конституенти об'єднати знаками диз'юнкції. Отримаємо ДДНФ: $f(x, y, z) = \neg x\neg y\neg z \vee \neg xy\neg z \vee x\neg y\neg z \vee x\neg yz$. Двоїсте правило існує для побудови ДКНФ. Розглянемо всі набори, на яких булева функція дорівнює нулю. Утворимо всі конституенти нуля як елементарні диз'юнкції, в яких кожна змінна береться без заперечення, якщо вона дорівнює нулю, та з запереченням, якщо вона дорівнює одиниці на даному наборі. З'єднаємо всі конституенти нуля символами кон'юнкції, отримаємо ДКНФ:

$$f(x, y, z) = (x \vee y \vee \neg z)(x \vee \neg y \vee \neg z)(\neg x \vee \neg y \vee z)(\neg x \vee \neg y \vee \neg z).$$

Інший спосіб полягає в тому, що будь-яку булеву формулу можна привести до вигляду ДНФ (КНФ) та ДДНФ (ДКНФ) за допомогою еквівалентних перетворень, використовуючи співвідношення (1.1) – (1.20).

Приклад. Приведемо до ДДНФ формулу $(x \oplus y)\rightarrow xz$. Отримаємо:

$(x\oplus y)\rightarrow xz = \neg(x\oplus y)\vee xz = (x\equiv y) \vee xz=\neg x\neg y\vee xy\vee xz$. Отримана формула представлена в ДНФ, і для того, щоб отримати її представлення в ДДНФ, необхідно в кожному «доданку» диз'юнкції (в кожному кон'юнкті) присутність всіх трьох змінних, від яких залежить функція. Отже, необхідно «помножити» кожний «доданок» на 1, утворену як диз'юнкцію змінної, якої не вистачає в даному кон'юнкті, і її заперечення. Таким чином, $\neg x\neg y$ та xy «помножуємо» на $1=z\vee\neg z$, та xz – на $1=y\vee\neg y$. Тоді отримаємо: $\neg x\neg y(z \vee \neg z) \vee xy(z \vee \neg z) \vee xz(y \vee \neg y) = xyz \vee xy\neg z \vee \neg x\neg yz \vee \neg x\neg y\neg z \vee$

xyz ∨ x¬yz. В силу ідемпотентності xyz ∨ xyz = xyz, тому ДДНФ є xyz ∨ xy¬z ∨ ¬x¬yz ∨ ¬x¬y¬z ∨ x¬yz. Аналогічно можна отримати ДКНФ.

1.5. Алгебра Жегалкіна

Визначення 1.12. Система елементів {0,1}, на якій визначені операції $\wedge$ (кон'юнкція) і $\oplus$ (сума за *mod* 2), для яких виконуються співвідношення:

(1.21) $x \oplus y = y \oplus x$,

(1.22) $x\,(y \oplus z) = xy \oplus xz$,

(1.23) $x \oplus x = 0$,

(1.24) $x \oplus 0 = x$,

а також співвідношення булевої алгебри (1.3, 1.6, 1.10, 1.11 - $x \wedge 1 = x$, $x \wedge 0 = 0$, $x \wedge (y \wedge z) = (x \wedge y) \wedge z$, $x \wedge y = y \wedge x$ і ін.), що відносяться до кон'юнкції і констант, називається *алгеброю Жегалкіна*.

З таблиці істинності операції суми за mod 2 слідує, що

(1.24) $\neg x = x \oplus 1$.

Операцію диз'юнкції можна виразити через $\oplus$ і $\wedge$ так:

(1.25) $x \vee y = \neg(\neg x\, \neg y) = (x \oplus 1)(y \oplus 1) \oplus 1 = xy \oplus x \oplus y$.

Визначення 1.13. Будь-яка формула алгебри Жегалкіна, що має вигляд суми (за mod 2) кон'юнкції булевих змінних, називається *поліномом Жегалкіна*.

Якщо в кожен член полінома Жегалкіна кожна змінна входить один раз і поліном не містить однакових членів, то такий поліном Жегалкина називається *канонічним*.

Теорема 1.6. Будь-яка булева функція може бути єдиним способом предсталена у вигляді канонічного поліному Жегалкіна.

Доведення. Будь-яку булеву формулу можна представити у вигляді полінома Жегалкіна, використовуючи співвідношення (1.24), (1.25). З (1.25) випливає, що якщо дві функції f1 і f2 такі, що $f_1 \wedge f_2 = 0$, то $f_1 \vee f_2 = f_1 \oplus f_2$. З цього слідує правило для представлення булевої функції у вигляді полінома

Жегалкіна: в булевої формулі, що задана у вигляді ДДНФ, досить замінити знак $\vee$ на знак $\oplus$, представити заперечення змінних як $\neg x = x\oplus 1$, розкрити дужки згідно із законом дистрибутивності (1.22) і привести подібні члени згідно (1.23, 1.24).◊

Приклад. Зведемо до канонічного поліному Жегалкіна булеву функцію: $f(x, y, z) = x\neg yz \vee x\neg y\neg z \vee xyz$. Оскільки функція знаходиться в ДДНФ, замінимо символи $\vee$ на $\oplus$, отримаємо: $f(x, y, z) = x\neg yz \oplus x\neg y\neg z \oplus xyz = x(y \oplus 1)z \oplus x (y \oplus 1) (z \oplus 1) \oplus xyz = xyz \oplus xz \oplus xyz \oplus xz \oplus xy \oplus x \oplus xyz = xyz \oplus xy \oplus x$.

Привести функцію до канонічного поліному Жегалкіна (КПЖ) можна і через суперпозицію відповідних поліномів функцій, які складають вихідну.

На приклад, для $x\rightarrow y$ ДДНФ є $\neg x\neg y \vee \neg xy \vee xy$. Після перетворення ми отримаємо КПЖ $xy \oplus x \oplus 1$.

З іншої сторони, $x \rightarrow y = \neg x \vee y=(x\oplus 1) y \oplus (x \oplus 1) \oplus y = xy \oplus x \oplus 1$. Або $x \rightarrow y = \neg(x\neg y)== 1 \oplus x (1 \oplus y) = xy \oplus x \oplus 1$.

Тоді для функції з попереднього прикладу $(x \oplus y)\rightarrow xz$ по ДДНФ $xyz \vee xy\neg z \vee \neg x\neg yz \vee \neg x\neg y\neg z \vee x\neg yz$ побудуємо КПЖ $xyz \oplus xz \oplus x \oplus y \oplus 1$. Або, замінивши імплікацію на відповідний КПЖ ($\oplus$ і & допустимі операції в алгебрі Жегалкіна), отримаємо $(x \oplus y)\rightarrow xz = (x \oplus y)xz \oplus (x \oplus y) \oplus 1 = xyz \oplus xz \oplus x \oplus y \oplus 1$.

Існують і інші способи побудови КПЖ: метод трикутника (трикутник Паскаля), перетворення Мебіуса.

1.6. Властивості булевих функцій

Визначення 9.14. Функція, яка виражена поліномом Жегалкіна вигляду $\sum\alpha_i x_i \oplus \gamma$, де α_i, γ є 0 або 1, називається *лінійною (L)*.

Всі функції однієї змінної лінійні. Лінійними функціями двох змінних є $x\oplus y$ і $x \equiv y$: $x \equiv y = \neg x\neg y \vee xy = (x\oplus 1)(y \oplus 1)\oplus xy = xy \oplus x\oplus y \oplus 1\oplus xy = x\oplus y \oplus 1$.

Визначення 1.15. Нехай дана деяка булева формула А. Формула А* називається *двоїстою* формулі А, якщо вона виходить з А шляхом заміни операцій кон'юнкції на диз'юнкцію, диз'юнкції на кон'юнкцію, 1 на 0, 0 на 1 усюди, де вони входять.

Наприклад: $A = xy \vee \neg xy$; $A^* = (x \vee y)(\neg x \vee y)$.

Теорема 1.7. Якщо $A(x_1,...,x_n)$ и $A^*(x_1,...,x_n)$ – дві взаємнодвоїсті формули, то $\neg A(x_1, ..., x_n) = A^*(\neg x_1, ..., \neg x_n)$.

Доведення випливає з законів де Моргана.

Визначення 1.16. Функція називається *самодвоїстою (S)*, якщо вона двоїста сама собі, тобто $A(x_1, x_2, ..., x_n) = \neg A(\neg x_1, \neg x_2, ..., \neg x_n)$, або $\neg A(x_1, x_2, ..., x_n) = A(\neg x_1, \neg x_2, ..., \neg x_n)$.

Для самодвоїстої функції заперечення її аргументів призводить до заперечення самої функції, отже, самодвоїста функція на протилежних наборах приймає протилежні значення. Наприклад, функція $xy \vee xz \vee yz$ самодвоїста. Щоб показати це, візьмемо заперечення від кожної змінної і від всієї функції::

$$\neg(\neg x\neg y \vee \neg x\neg z \vee \neg y\neg z) = (x\vee y)(x\vee z)(y\vee z) = (x \vee xy \vee xz \vee yz)(y\vee z) =$$

$$= xy\vee xy \vee xyz \vee yz\vee xz \vee xyz\vee xz \vee yz = xy \vee xz \vee yz.$$

Відношення порядку на наборах булевих змінних визначається так. Розглянемо два набори: $A = (\alpha_1,..., \alpha_i, ..., \alpha_n)$ і $B = (\beta_1,..., \beta_i,..., \beta_n)$. Якщо для всіх i ($i = 1, ..., n$) $\alpha_i \leq \beta_i$, і існує хоча б одне таке j, при якому $\alpha_j < \beta_j$, набір А передує (менше) набору В. Це позначається: $A \leq B$. Якщо для деяких i $\alpha_i \leq \beta_i$ і існує таке j, що $\alpha_j > \beta_j$, то набори A і B *непорівнювані*. Наприклад: набори (0101) і (0010) непорівнювані, а набір (0101) передує набору (0111).

Тоді булеву функцію можна задати на цій решітці, як показано на рис. 1.1. Причому серед булевих функцій можна виділити такі, які не зменшуються зі зростанням значень порівнянних наборів на булевої решітці. Такі функції називаються монотонними. Наприклад, на рис. 1.1. показана монотонна функція (00001111).

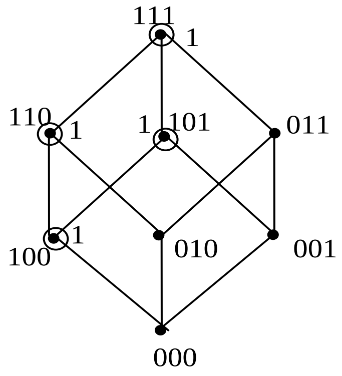


Рис. 1.1. Монотонна функція.

Визначення 1.17. Булева функція f називається *монотонною (M)*, якщо для будь-яких двох наборів A і B з її області визначення таких, що $A \le B$, $f(A) \le f(B)$. Якщо хоча б для однієї пари наборів таких, що $A \le B$, $f(A) > f(B)$, то функція немонотонна.

Визначення 1.18. Булева функція називається *функцією, що зберігає нуль* (T_0), якщо на нульовому наборі вона дорівнює нулю, тобто $f(0,\ldots,0) = 0$.

Визначення 1.19. Булева функція називається *функцією, що зберігає одиницю* (T_1), якщо на одиничному наборі вона дорівнює одиниці, тобто $f(1,\ldots,1) = 1$.

Властивості основних функцій наведено в табл.1.6.

1.7. Замкнуті класи булевих функцій

Нехай функція f(x) представлена формулою F(x), а функція g представлена формулою G. Тоді формула F(G), яка виходить шляхом заміни кожного входження змінної x в формулі F(x) на формулу G, представляє суперпозицію булевих функцій f (g).

Формула $F = \overline{(x \vee y)}z$ представляє суперпозицію функцій $F_1(F_3(F_7(x,y),z))$, де F_1-кон'юнкція, F_7-диз'юнкція, F_3-заперечення. Суперпозицію функцій $F_6(F_1(F_{13}(x,y),F_7(x,y)),z)$ можна записати у вигляді формули $((x \to y) \wedge (x \vee y)) \oplus z$, тут F_6 - складання mod 2, F_{13} - імплікація.

Визначення 1.20. Клас функцій називається *функціонально замкнутим*, якщо суперпозиція цих функцій належить даному класу.

Теорема 1.8. Суперпозиція лінійних функцій є функція лінійна.

Доведення. Якщо в лінійний поліном Жегалкіна на місце будь-якої змінної підставити лінійну функцію, то отриманий поліном буде також лінійним. Візьмемо поліном: $\alpha_1 x_1 \oplus \ldots \oplus \alpha_i x_i \oplus \ldots \oplus \alpha_n x_n \oplus \gamma$. Підставимо замість x_1 лінійний поліном $\Sigma\beta_i y_i \oplus \xi$.

Отримаємо лінійний поліном: $\alpha_1\beta_1 y_1 \oplus \ldots \oplus \alpha_1\beta_i y_i \oplus \ldots \oplus \alpha_1\beta_m y_m \oplus \alpha_1\xi \oplus \ldots \oplus \alpha_n x_n \oplus \gamma$. Отже, клас лінійних функцій функціонально замкнутий. ◊

Таблиця 1.6.

		КПЖ	T_0	T_1	M	S	L
f_0 –	константа 0	0	+	-	+	-	+
f_1 –	$x \wedge y$ – кон'юнкція	xy	+	+	+	-	-
f_3 –	тотожна функція $f(x,y) = x$	x	+	+	+	+	+
f_6 –	$x \oplus y$ – сума за mod 2 (операція Жегалкіна)	$x \oplus y$	+	-	-	-	+
f_7 –	$x \vee y$ – диз'юнкція	$xy \oplus x \oplus y$	+	+	+	-	-
f_8 –	$x \downarrow y$ – стрілка Пірса	$xy \oplus x \oplus y \oplus 1$	-	-	-	-	-
f_9 –	$x \equiv y$ (або $x \sim y$) – еквівалентність	$x \oplus y \oplus 1$	-	+	-	-	+
f_{12} –	заперечення $f(x) = \neg x$	$x \oplus 1$	-	-	-	+	+
f_{13} –	$x \to y$ – імплікація	$xy \oplus x \oplus 1$	-	+	-	-	-
f_{14} –	$x \mid y$ – штрих Шеффера	$xy \oplus 1$	-	-	-	-	-
f_{15} –	константа 1	1	-	+	+	-	+

Теорема 1.9. Суперпозиція монотонних функцій є функція монотонна. Отже, клас монотонних функцій є функціонально замкнутим.

Доведення. Нехай функції $f(x_1,\ldots, x_n)$, $g_1(y_1, \ldots, y_k)$, …, $g_n(y_1, \ldots, y_k)$ монотонні. Складемо суперпозицію функцій $\varphi = f(g_1, \ldots, g_n)$. Нехай α і β – два набори значень змінних $y_1, \ldots, y_k$, причому $\alpha \le \beta$. В силу монотонності функцій $g_1, \ldots, g_n$

$g_1(\alpha) \le g_1(\beta), \ldots, g_n(\alpha) \le g_n(\beta)$,

тому набори значень функцій впорядковані:

$(g_1(\alpha), \ldots, g_n(\alpha)) \le (g_1(\beta), \ldots, g_n(\beta))$,

а в силу монотонності функції f

$f(g_1(\alpha), \ldots, g_n(\alpha)) \le f(g_1(\beta), \ldots, g_n(\beta))$.

Звідси отримуємо $\varphi(\alpha) \le \varphi(\beta)$. ◊

Теорема 1.10. Клас функцій, що зберігають нуль, є *функціонально замкнутим.*

Доведення. Нехай функції $f(x_1,\ldots, x_n)$, $g_1(y_1, \ldots, y_k)$, …, $g_n(y_1, \ldots, y_k)$ зберігають нуль. Складемо суперпозицію функцій $\varphi = f(g_1, \ldots, g_n)$. Тоді

$\varphi(0,\ldots,0) = f(g_1(0,\ldots,0), \ldots, g_n(0,\ldots,0)) = f(0,\ldots,0) = 0$. ◊

Теорема 1.11. Клас функцій, що зберігають одиницю, є функціонально замкнутим.

Доводиться аналогічно теоремі 1.10.

Теорема 1.12. Клас самодвоїстих функцій є функціонально замкнутим.

Доведення. Нехай функції $f(x_1,\ldots, x_n)$, $g_1(y_1, \ldots, y_k)$, …, $g_n(y_1, \ldots, y_k)$ самодвоїсті. Складемо суперпозицію функцій $\varphi = f(g_1, \ldots, g_n)$. Тоді

$\varphi^* = f^*(g_1^*, \ldots, g_n^*) = f(g_1, \ldots, g_n) = \varphi$. ◊

1.8. Функціональна повнота систем булевих функцій

Визначення 1.21. Система булевих функцій називається *функціонально повною,* якщо будь-яка булева функція може бути представлена як суперпозиція функцій з цієї системи.

Позначимо: T_0 – клас функцій, що зберігають 0; T_1 – клас функцій, що зберігають 1; S – клас самодвоїстих функцій; M – клас монотонних функцій; L – клас лінійних функцій.

Теорема 1.13 (Поста). Для того, щоб система функцій була повна, необхідно і достатньо, щоб вона містила хоча б одну немонотонну, хоча б одну нелінійну, хоча б одну несамодвоїсту, хоча б одну, що не зберігає нуль, і хоча б одну, не зберігаючу одиницю, функцію.

Доведення. Необхідність умови теореми виходить з функціональної замкнутості і неповноти класів монотонних, лінійних, зберігаючих 0, зберігаючих 1 і самодвоїстих функцій. Доведено (теореми 1.9 – 1.12), що функція, що не належить цьому функціонально замкнутому класу, не може бути побудована шляхом суперпозиції функцій цього класу.

Для доведення *достатності* покажемо, що за допомогою функцій, що не належать класам T_0, T_1, S, M, L можна побудувати деяку повну систему функцій. Такою (повною) системою є, наприклад, заперечення і кон'юнкція. Дійсно, будь-яка булева функція представлена у вигляді ДДНФ, тобто як суперпозиція $\neg, \wedge, \vee$. Отже, система $\{\neg, \wedge, \vee\}$ функціонально повна. Можна виключити з неї $\vee$, так вона представлена як суперпозиція $\neg$ і $\wedge$: $x \vee y =$ $=\neg(\neg x \wedge \neg y)$. Доведення носить конструктивний характер.

Спочатку побудуємо константи. Якщо функція $f(x_1,\ldots, x_n)$ несамодвоїста, то підстановкою в неї x і $\neg x$ можна отримати константу. Дійсно, зважаючи на несамодвоїстість $f(x_1,\ldots, x_n)$ існує такий набір $(\alpha_1,\ldots, \alpha_n)$, що $f(\alpha_1,\ldots, \alpha_n) =$ $f(\neg\alpha_1,\ldots, \neg\alpha_n)$. Тоді функція $\varphi(x) = f(x^{\alpha 1},\ldots, x_n^{\alpha n})$ є константою, так як $\varphi(0) =$ $f(0^{\alpha 1},\ldots, 0^{\alpha n}) = f((\neg\alpha_1, \ldots, \neg\alpha_n) = f(\alpha_1, \ldots, \alpha_n) = f(1^{\alpha 1},\ldots, 1^{\alpha n}) = \varphi(1)$.

Константу 1 можна побудувати за допомогою функції, що не зберігає 0 і зберігає 1, тобто несамодвоїстої. Дійсно, нехай f не зберігає 0. Тоді $\varphi(0) =$ $f(0,\ldots,0) = 1$, а якщо при цьому f зберігає 1, то $\varphi(1) = f(1,\ldots,1) = 1$, тобто константа 1 побудована.

Двоїсто, нехай f_1 не зберігає 1. Тоді $\psi(x) = f_1(\varphi(x), \ldots, \varphi(x)) = f_1(1,\ldots,1)) = 0$, тобто $\psi(x)$ існує константа 0.

Якщо ж $f(1, \ldots,1) = 0$, то $\varphi(x)=\neg x$, так як $\varphi(0)=1$ за визначенням f, а $\varphi(1)=0$. Тоді, якщо взяти несамодвоїсту функцію f_2 , то підставляючи в неї x і $\neg x$, так само можна отримати константу, а за допомогою ще одного заперечення, получити іншу константу. Таким чином, за допомогою функцій, тих, які не зберігають 0, не зберігають 1 і несамодвоїстих, можемо побудувати константи 0 та 1.

За допомогою немонотонної функції підстановкою в неї констант можемо отримати запереченння. Дійсно, нехай f немонотонна. Тоді існують набори α і β, такі, що $\alpha \le \beta$, $f(\alpha) = 1$, $f(\beta) = 0$. Оскільки $\alpha \le \beta$, то в α є декілька, наприклад, k-елементів, які рівні 0, в той час як в β ці ж елементи рівні 1. Виберемо набір α та замінемо у ньому перший такий нульовий елемент на 1, отримаємо набір $\alpha \le \alpha^1$, який відрізняється від α тільки одним елементом (такі набори називають сусідніми). Повторюючи цю операцію k разів, отримаємо послідовність наборів $\alpha \le \alpha^1 \le\ldots\le \alpha^{k-1} \le \beta$, в якій будь-які два сусідні набори відрізняються один від одного тільки одним елементом. В цьому ланцюжку знайдуться два такі набори α^i, α^{i+1}, що $f(\alpha^i) = 1$, $f(\alpha^{i+1}) = 0$. Нехай ці набори відрізняються i-м елементом (значенням змінної x_i), інші елементи однакові. Підставимо в f ці значення. Тоді отримаємо функцію $f(\alpha^i_1, \ldots, x_i, \alpha^i_{i+1}, \ldots, \alpha^i_n) = g(x_i)$, яка залежить тільки від x_i. Тоді $g(0) = g(\alpha^i_i) = f(\alpha^i) = 1$, $g(1) = g(\alpha^{i+1}_i) = f(\alpha^{i+1}) = 0$. Звідси слідує, що $g(x_i) = \neg x_i$.

За допомогою підстановки в нелінійну функцію констант і використання заперечення можемо побудувати кон'юнкцію і диз'юнкцію. Дійсно, якщо функція f нелінійна, то її поліном Жегалкіна містить кон'юнкцію змінних, наприклад, $K = x_1x_2\ldots x_k$. Нехай $x_3 = \ldots = x_k = 1$, а для всіх x_j, тих, що не містяться в K, $x_j = 0$. Тоді кон'юнкція $K = x_1x_2$, інші кон'юнкції перетворяться в 0, а поліном Жегалкіна стане виду $\varphi(x_1, x_2) = x_1x_2 \oplus \alpha_1x_1 \oplus \alpha_2x_2 \oplus \gamma$, де α_1, α_2, γ – коефіцієнти, равні 0 або 1. При підстановці різних констант замість α_1, α_2, γ

будуть отримані різні функції. Розглянемо функцію $\psi(x_1, x_2)$, отриману з $\varphi(x_1, x_2)$ наступним чином:

$\psi(x_1, x_2) = \varphi(x_1 \oplus \alpha_2, x_2 \oplus \alpha_1) \oplus \alpha_1\alpha_2 \oplus \gamma = (x_1 \oplus \alpha_2)(x_2 \oplus \alpha_1) \oplus \alpha_1(x_1 \oplus \alpha_2) \oplus \alpha_2(x_2 \oplus \alpha_1) \oplus \gamma \oplus \alpha_1\alpha_2 \oplus \gamma = x_1x_2$.

Таким чином, кон'юнкція отримана.

Для отримання диз'юнкції по законам де Моргана потрібно тільки заперечення. ◊

Приклад. Системи функцій $\{\vee, \wedge, \neg\}$, $\{\oplus, \wedge, 0, 1\}$, $\{\rightarrow, \neg\}$ являються функціонально повними. Для прикладу перевіримо повноту системи функцій $\{\neg, \rightarrow\}$. Для досліджуваної системи складемо таблицю Поста – критерійну таблицю (табл.1.7). Якщо функція входить у функціонально замкнутий клас, то в таблиці Поста у відповідній чарунці ставиться знак '+', інакше – знак '–'.

Функція $\neg x$ не зберігає 0 і 1, так як на нульовому наборі вона набуває значення 1, а на одиничному - 0. Очевидно, що ця функція немонотонна. Функція самодвоїста, так як на протилежних наборах вона набуває протилежних значень. Функція лінійна – її поліном Жегалкіна: $\neg x = x \oplus 1$. Функція $x \rightarrow y$ не зберігає 0 і зберігає 1. Ця функція немонотонна, так як набір (0,0) передує набору (1,0), но $0 \rightarrow 0 = 1$, а $1 \rightarrow 0 = 0$. На протилежних наборах (0, 0) і (1, 1) функція набуває однакових значень 1, отже, вона несамодвоїста.

Для перевірки лінійності $x \rightarrow y$ побудуємо канонічний поліном Жегалкіна:

$x \rightarrow y = \neg x \neg y \vee \neg xy \vee xy = (x \oplus 1)(y \oplus 1) \oplus (x \oplus 1)y \oplus xy = xy \oplus x \oplus 1$. Функція нелінійна, так як містить елемент xy.

Таблиця 1.7

	T_0	T_1	S	M	L
$\neg$	–	–	+	–	+
$\rightarrow$	–	+	–	–	–

Система функцій {¬, →} повна, так як в кожному стовпці таблиці Поста є хоч би один знак '–'.

Розрізняють повноту в слабкому і сильному сенсі. Повнота в слабкому сенсі припускає використання функціонально повних констант 0 і 1. Якщо ж унеможливити використання констант, то система функцій повна в сильному сенсі.

Повну систему називають неприводимою або базисом, якщо з неї не можна виключити жодної функції без того, щоб система не втратила властивості повноти.

Теорема 1.14. Будь-яка неприводима функціональна система булевих функцій складається з не більш ніж чотирьох функцій. Над доведенням пропонується подумати самостійно.

Функціонально повна система називається незалежною, якщо жодна з функцій системи не може бути суперпозицією інших функцій цієї системи. Базис – незалежний. Наприклад система {⊕, ∧, 0, 1} є повною, але не є незалежною. Базисом буде система {⊕, ∧, 1}, так як $0=1\oplus 1$ (табл.1.8).

Таблиця 1.8

	T_0	T_1	S	M	L
⊕	+	–	–	–	+
∧	+	+	–	+	–
1	–	+	–	+	+
0	+	–	–	+	+

Тезу про повноту можна узагальнити. Якщо усі функції функціонально повної системи можна виразити через супрепозицію функцій іншої системи, то вона так само буде функціонально повною. Цей метод називається методом зведення. Наприклад, система алгебри Жегалкіна {⊕, ∧, 1} є функціонально повною. Система {≡, ∨, 0} також буде функціонально повною. Дійсно, можно виразити $x \oplus y = (x \equiv y) \equiv 0$, $x \wedge y = (x \vee y) \equiv (x \equiv y)$, $1 = 0 \equiv 0$**.** Та переконатися**,** побудувавши критеріальну таблицю (табл.1.9):

Таблиця 1.9

	T_0	T_1	S	M	L
$\equiv$	–	+	–	–	+
$\vee$	+	+	–	+	–
0	+	–	–	+	+

1.9. Мінімізація булевих функцій

Визначення 1.22. *Імплікантою* булевої функції $f(x_1, ..., x_n)$ називається така булева функція $g(x_1, ..., x_n)$, яка дорівнює одиниці на деяких з тих наборів, на яких дорівнює одиниці дана функція, і дорівнює нулю на інших наборах (із визначения випливає, що функція g – *імпліканта* f, якщо $g \to f \equiv 1$.) Говориться, що імпліканта g покриває своїми одиницями деякі одиниці даної функції f.

Якщо функція задана у вигляді ДНФ, тобто у вигляді диз'юнкції елементарних кон'юнкцій, то кожний кон'юнкт являється імплікантою даної функції. Довжина деяких елементарних кон'юнкцій може бути зменшена за допомогою еквівалентних перетворень. Для цього застосовуються такі співвідношення:

(1.26) *закон неповного склеювання*:

$xy \vee \neg xy = y$;

$(x \vee y)(\neg x \vee y) = y$;

(1.27) *закон повного склеювання*:

$xy \vee \neg xz = xy \vee \neg xz \vee yz$;

$(x \vee y)(\neg x \vee z) = (x \vee y)(\neg x \vee z)(y \vee z)$;

(1.28) *закони поглинання*:

$x \vee xy = x$;

$x(x \vee y) = x$;

Доведення законів повного склеювання:

$\boldsymbol{xy \vee \neg xz \vee yz} =$

$= xy \vee \neg xz \vee yz(x \vee \neg x) =$

$= xy \vee \neg xz \vee xyz \vee \neg xyz =$

$xy \vee xyz \vee \neg xz \vee \neg xyz =$

$= \boldsymbol{xy \vee \neg xz}$.

(1.29) $x \vee \neg xy = x \vee y$;

$x(\neg x \vee y) = xy$.

$\boldsymbol{xy \vee \neg xz} =$

$=(x \vee \neg x)(x \vee z)(y \vee \neg x)(y \vee z)$

$= 1(xy \vee x\neg x \vee zy \vee z\neg x) =$

$= xy \vee 0 \vee zy \vee z\neg x =$

$\boldsymbol{xy \vee \neg xz \vee yz}$.

Визначення 1.23. Елементарна кон'юнкція, яка є імплікантою функції f, але жодна її власна частина імплікантою цієї функції не являється, називається *простою імплікантою* даної функції.

Довжина простої імпліканти вже не може бути зменшена шляхом склеювання її з іншими імплікантами даної функції.

Визначення 1.24. Диз'юнкція всіх простих імплікант булевої функції називається *скороченою диз'юнктивною нормальною формою* булевої функції.

Приклад. Скоротимо формулу: $F(x, y, z) = xyz \vee x\neg yz \vee xy\neg z \vee \neg xy\neg z$. Склеїмо конституенти: $xyz \vee x\neg yz = xz$, $xy\neg z \vee \neg xy\neg z = y\neg z$, $xyz \vee xy\neg z = xy$. Отримаємо скорочену форму: $F(x, y, z) = xz \vee y\neg z \vee xy$.

Множина усіх простих імплікант покриває усі одиниці булевої функції. Але представлення булевої функції у вигляді скороченої ДНФ ще не являється найбільш економним. Серед множини простих імплікант можуть опинитись зайві, тобто такі, що покривають одиниці функції, вже покриті іншими імплікантами. Видаляючи зайві імпліканти, можна отримати *тупикову* диз'юнктивну нормальну форму. Тупикова ДНФ, яка містить найменшу кількість імплікант та змінних, називается *мінімальною* ДНФ.

За допомогою елментів можна побудувати логічну схему з n входами та одним виходом, реалізуючу будь-яку булеву функцію n змінних. Якщо на усі входи цієї схеми подати сигнали, рівні нулю або одиниці, то на виході з'явиться сигнал, відповідний до значення функції на даному наборі (0 або 1).

Синтез будь-якої логічної схеми повністю аналогічний побудові формули, що визначає булеву функцію.

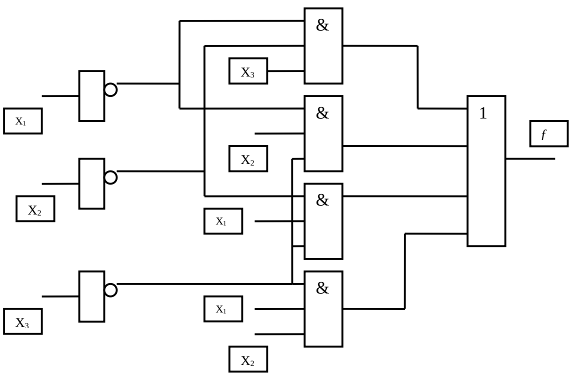
Нехай булева функція $f(x_1, x_2, x_3)$ задана своєю таблицею істинності :

X1	0	0	0	0	1	1	1	1
X2	0	0	1	1	0	0	1	1
X3	0	1	0	1	0	1	0	1
f	0	1	1	0	1	0	1	0

Потрібно побудувати логічну схему, описану даною функцією.

Представимо функцію $f(x_1, x_2, x_3)$ у вигляді ДДНФ:

$f = \overline{x}_1\overline{x}_2 x_3 \vee \overline{x}_1 x_2 \overline{x}_3 \vee x_1 \overline{x}_2 \overline{x}_3 \vee x_1 x_2 \overline{x}_3$ Логічна схема, реалізуюча дану формулу, може бути побудована на логічних елементах «або», «і» та інверторах. Побудуємо таку схему.



Ця схема містить 3 інвертори, 4 елемента «і» з трьома входами, 1 елемент типу «або» з чотирма входами.

Виникає питання, чи не можна перетворити формулу таким чином, щоб зменшити кількість логічних елементів в схемі і кількість входів цих елементів. Очевидно, тоді схема стане простіше і, при її технічній реалізації, дешевше. Спробуємо це зробити. В останніх двох членах ДДНФ функції f винесемо за дужку $x_1\overline{x}_3$.

Отримаємо $f_1 = \overline{x_1}\,\overline{x_2}x_3 \vee \overline{x_1}x_2\overline{x_3} \vee x_1\overline{x_3}(\overline{x_2} \vee x_2) = \overline{x_1}\,\overline{x_2}x_3 \vee \overline{x_1}x_2\overline{x_3} \vee x_1\overline{x_3}$. Так як $x_2 \vee \overline{x_2} = 1$, то $f_1 = f$.

Для реалізації формули f_1 знадобиться 3 інвертори, 3 елемента «і» із спільною кількістю входів, рівною 8, і один елемент «або» з 3 входами.

В залежності від поставленої мети мінімізації, вибір мінімальної ДНФ можливий за двома критеріями:

- вміст найменшої кількості входження букв - загальна кількість букв у всіх елементарних кон'юнкціях, так звана ціна покриття C^{α} *(для нашого прикладу в початковій формулі була 12, стало 8);*
- сума кількості входжень букв та кількості елементарних добутків являється найменшою (ціна покриття C^{β} - *для нашого прикладу була 20, стало 12).*

В цих наведених прикладах при мінімізації застосовувався метод Блейка, оснований на законах склеювания та поглинання.

1.9.1. Метод Квайна отримання скороченої диз'юнктивної нормальної форми

Теорема 1.15 (Квайна). Якщо в досконалій диз'юнктивній нормальній формі булевої функції зробити всі операції неповного склеювання (1.26), а потім всі операції поглинання (1.27, 1.28), то в результаті буде отримана скорочена ДНФ даної функції.

Для перетворення довільної ДНФ до виду ДДНФ необхідно застосувати до елементарних кон'юнкцій, що містять не всі змінні, операцію розгортання, наприклад: $xy(z \vee \neg z) = xyz \vee xy\neg z$, де z – недостатня змінна.

Приклад. Булеву функцію задано у вигляді: $f(x,y,z) = \neg xyz \vee x\neg y\neg z \vee \neg yz$. Приведемо формулу до ДДНФ:

$f(x, y, z) = \neg xyz \vee x\neg y\neg z \vee \neg yz(x \vee \neg x) = \neg xyz \vee x\neg y\neg z \vee x\neg yz \vee \neg x\neg yz$.

Випишемо всі конституенти одиниці і зробимо всі операції неповного склеювання:

* $\neg xyz$	$\neg yz$
* $x\neg yz$	$x\neg y$
* $x\neg y\neg z$	$\neg xz$
* $\neg x\neg yz$	

Відзначимо всі кон'юнкції, що склеюються символом «*». Вони поглинаються імплікантами, отриманими в результаті склеювання. Невідмічені кон'юнкції нічим не поглинаються і є простими імплікантами.

Скорочена ДНФ даної функції: $f(x,y,z)= \neg yz \vee x\neg y \vee \neg xz$. Для знаходження мінімальної ДНФ складемо імплікантну матрицю (таблиця 1.10), у якій по рядкам записуємо імпліканти, по стовпцям – конституенти одиниці.

Таблиця 1.10

	$\neg xyz$	$x \neg y \neg$	$x \neg yz$	$\neg x \neg yz$
$\neg yz$			×	×
$x\neg y$ *		×	×	
$\neg xz$ *	×			×
	*	*	+	+

У клітинах таблиці хрестиками відзначаємо імпліканти, що покривають одиниці даної функції. Внизу таблиці символом «*» відзначаємо ті стовпці, в яких стоїть тільки один хрестик, відповідні їм імпліканти також відзначаємо символом «*» – вони є обов'язковими. Відзначаємо також символом «+» ті стовпці, які покриваються обов'язковими імплікантами. Якщо всі стовпчики відзначені, то отриманий набір обов'язкових імплікант становить мінімальну ДНФ. Якщо частина стовпців залишається непокритою, з решти імплікант вибирається найменше число найбільш коротких імплікант так, щоб всі стовпці були покриті.

У нашому прикладі мінімальна ДНФ: $f(x, y, z) = x\neg y \vee \neg xz$.

Розглянемо фунцію (01111110). Імплікативна матриця буде наступною:

		$\bar{x}y\bar{z}$	$\bar{x}yz$	$x\bar{y}z$	$\bar{x}\bar{y}z$	$xy\bar{z}$	$x\bar{y}\bar{z}$
A	$x\bar{z}$					X	X
B	$y\bar{z}$	X				X	
C	$\bar{y}z$			X	X		
D	$x\bar{y}$			X			X
E	$\bar{x}y$	X	X				
F	$\bar{x}z$		X		X		

Обов'язкових імплікант немає. Серед необов'язкових можна вибрати тупикові ДНФ:

$$\varphi = y\bar{z} \vee x\bar{y} \vee \bar{x}z$$
$$\varphi = x\bar{z} \vee \bar{y}z \vee \bar{x}y$$

Ці форми мають однакову ціну покриття; отже, обидві є мінімальними.

Для знаходження всіх тупикових форм функції можна застосувати наступний прийом. Позначимо кожну імпліканту буквою: A, B, C ... . Поставимо у відповідність імпліканте значення 1, якщо імпліканта входить в покриття, і 0, якщо ні. Для покриття кожної одиниці необхідно, щоб в покриття входила хоча б одна імпліканта, відповідна позначці в даному стовпці. Так, для покриття першого стовпчика необхідно, щоб в тупикову ДНФ функції увійшла або імпліканта $y\bar{z}$, або $\bar{x}y$, тобто повинно бути $B \vee E = 1$. Це умова покриття першого стовпця матриці. Умова покриття всіх одиниць функції буде виглядати так:

$$(B \vee E)(E \vee F)(C \vee D)(C \vee F)(A \vee B)(A \vee D) = 1,$$

або, після перетворень: $ACE \vee BDF \vee ABCF \vee ADEF = 1$.

Для покриття необхідно вибрати систему з імплікант A, C і E; або B, D і F. Це рівнозначні варіанти, так як ціна покриття у них однакова. Ще два варіанти мають ціну покриття більшу.

Цей метод можна використовувати для знаходження базису функціонально повної приводимої системи функцій. Нехай у нас є таблиця Поста для системи з чотирьох функцій {A, B, C, D}.

	T_0	T_1	S	M	L
A	+	–	–	+	–
B	+	+	–	+	–
C	–	–	+	–	+
D	–	+	–	–	+

Тоді умови повноти можна записати наступним чином (по «мінусах» в кожному стовпці):

$(C\vee D)(A\vee D)(A\vee B\vee D)(C\vee D)(A\vee B)=(A\vee D)(A\vee B\vee D)(C\vee D)(A\vee B)=$

$=(A\vee D)(A\vee B\vee D)(C\vee D)(A\vee B)=(A\vee D)(A\vee B)(C\vee D)=(A\vee BD)(C\vee D)=AC\vee AD\vee$

$\vee BCD\vee BDD= AC\vee AD\vee BCD\vee BD =AC\vee AD\vee BC.$

Функціонально повними неприведеними системами є {A,C} або {A,D} або {B,C}.

Метод Квайна містить незручності, пов'язані з необхідністю попарного порівняння імплікант на етапі визначення простих імплікант. Із зростанням кількості мінтерм в СДНФ збільшується кількість цих порівнянь, що характеризується факторіальною функцією.

У 1956 році Маккласки запропонував модернізацію першого етапу методу Квайна. Ідея полягає в тому, що, якщо записати мінтерми у вигляді двійкових чисел і розбити їх на непересічні групи за кількістю 1 в двійковому числі, то попарно порівнювати потрібно тільки елементи сусідніх груп. Для попереднього прикладу ми отримаємо 001, 010, 011, 100, 101, 110 і розіб'ємо на групи (001, 010, 100) і (011, 101, 110). В результаті порівняння ми отримаємо 0-1, -01, 01-, -10, 10, 1-0, що і відповідає $\neg xz$, $\neg yz$, $\neg xy$, $y\neg z$, $x\neg y$, $x\neg z$.

На підставі отриманого будується імплікантна матриця.

1.9.2. Мінімізація булевих функцій за допомогою діаграм Вейча

Булеві функції можуть бути представлені графічно у вигляді діаграм Вейча або карт Карно (інколи зустрічається інша транскрипція вимови – Карнау) – вони відрізняються лише позначеннями клітинок таблиці.

Діаграма Вейча представляє собою прямокутник, для однієї змінної – поділений навпіл (одна половина відповіє змінній, друга - її запереченню). При введені кожної нової змінної відбувається повторне ділення діаграми навпіл.

Розглянемо діаграми Вейча для функцій від трьох і чотирьох змінних (рис. 1.2).

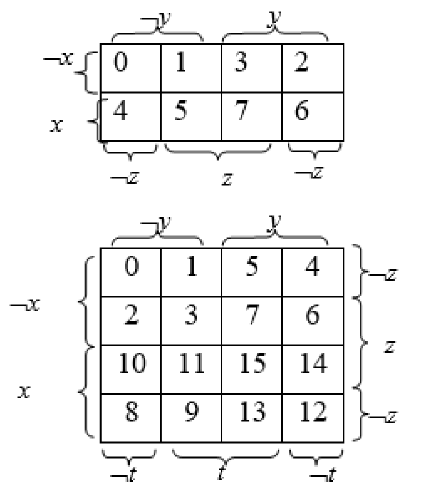


Рис. 1.2. Діаграми Вейча

Кожна клітинка діаграми відповідає одному набору змінних, номер клітинки – це двійковий код набору. При заданні булевої функції у відповідну до номера набору клітинку записується одиниця, і, таким чином, кожна клітинка діаграми з 1 представляє собою одну конституенту одиниці.

Приклад. Нехай функція $f(x, y, z, t) = 1$ на наборах 0, 1, 2, 3, 6, 7, 8, 15. Діаграма Вейча для заданої функції зображена на рис. 1.3. Одиниці функції, записані в сусідніх клітинках, відрізняються значенням тільки однієї змінної, внаслідок, вони склеюються по цій змінній і утворюють імпліканту. У цьому випадку кажуть, що імпліканта *покриває* відповідні одиниці булевої функції.

Наприклад, дві одиниці на наборах 7 і 15, покриваються імплікантою *yzt*, чотири одиниці на наборах 2,3,7,6 – імплікантою ¬*xz*. При цьому сусідніми вважаються також клітинки, записані вздовж лівого та правого краю діаграми (відрізняються значенням *y*) і вздовж верхнього і нижнього краю (відрізняються значенням *x*). Тобто, діаграму можно вважати ніби згорнутою в тор (а для трьох змінних – у циліндр).

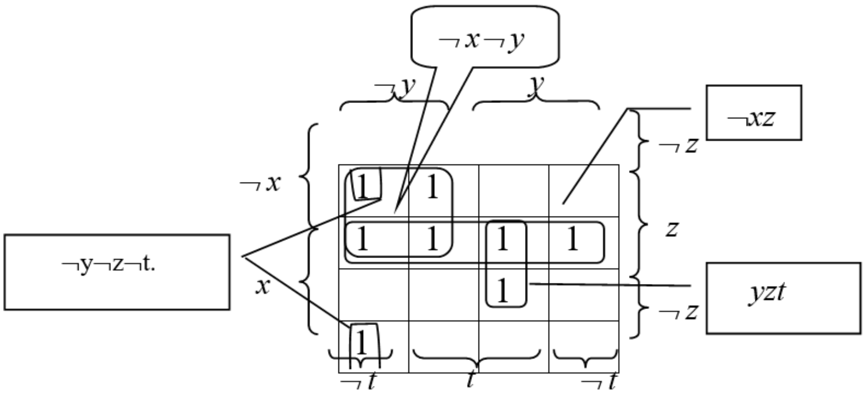


Рис. 1.3. Мінімальна ДНФ

При мінімізації булевої функції на діаграмі Вейча спочатку знаходять покриття, які містять максимальне число одиниць (8, 4, 2), а потім покриття, яке покриває решту одинить таким чином, щоб вони також були максимальні по величині, і при видаленні цього покриття хоча б одна одиниця функції залишилася непокритою. При цьому деякі одиниці можуть бути покриті неодноразово. Для функції, зображеної на рис. 1.3, мінімальна ДНФ: $\neg x\neg y \vee \neg xz \vee yzt \vee \neg y\neg z\neg t$.

Мінімальна КНФ будується двоїсто за діаграмою Вейча, заповненої нулями в порожніх клітинках (для збереження номерів наборів при підписуванні частин діаграми змінні та заперечення міняються місцями, тому що при побудові ДКНФ заперечення ставиться над 1 в значеннях набору, а в ДДНФ – над 0). Для даної функції мінімальна КНФ: $(\neg y\vee z)(\neg x\vee\neg z\vee t)(\neg x\vee y\vee\neg t)$.

Не повністю задана функція наноситься на діаграму Вейча таким самим чином, як і повністю задана, але клітинки, де функція не визначена, відрізняються символом *.

Мінімізація неповної функції відбувається наступним чином. Для отримання простих імплікант відбувається склеювання так, як якби в клітинках, відмічених *, стояли б одиниці. Якщо ж деякі клітинки, відмічені символом *, не попадають ні в одне покриття, вважаємо, що в них стоять нулі.

Карти Карно також служать для графічного представлення булевих функцій і відрізняються від діаграм Вейча тільки формою представлення змінних. Рядки і стовпці карти Карно відмічені наборами змінних, відповідних даному рядку або стовпцю.

Карта Карно для трьох змінних. На ній відмічена функція $\neg x\ \neg yz \vee x\ y \neg z$.

yz x	00	01	11	10
0		1		
1				1

Карта Карно для чотирьох змінних. Мінімальна форма відміченої функції $\neg y \neg t \vee \neg xz \neg t \vee xyzt$.

yz xt	00	01	11	10
00	1	1	1	
01				
11			1	
10	1	1		

Питання для самоконтролю до розділу 1

1. Визначення булевої алгебри, функції. Функції від 1 та 2-х зміних, іх кількість.
2. Теоретичні заснови побудови ДДНФ та ДКНФ.
3. Канонічний поліном Жегалкіна.

4. Перелічити та дати визначення властивостям булевих функцій.

5. Функціонально замкнені класи булевих функцій.

6. Функцональна повнота булевих функцій. Теорема Поста.

7. Теорема о максимальній кількості булевих функцій в функціонально повній системі.

8. Довести функціональну повноту системи, що складається з однієї функції – штрих Шеффера/стрілка Пірса.

9. Перелічити методи мінімізації булевих функцій. Цена покриття.

10. В чому суть мінімізації за діаграмою Вейча.

2. АЛГЕБРА ВИСЛОВЛЮВАНЬ

2.1. Методологічні принципи формальної логіки

Логіка - наука про мислення, про правильні міркування, про прийоми та методи пізнання за допомогою міркувань. В кінці 19-го – початку 20-го століття з появою булевої алгебри почався бурхливий розвиток математичної (формальної) логіки. На даний час розрізняють традиційну логіку як частину філософії, і формальну, математичну логіку, співвідношення яких показано на рис. 2.1.

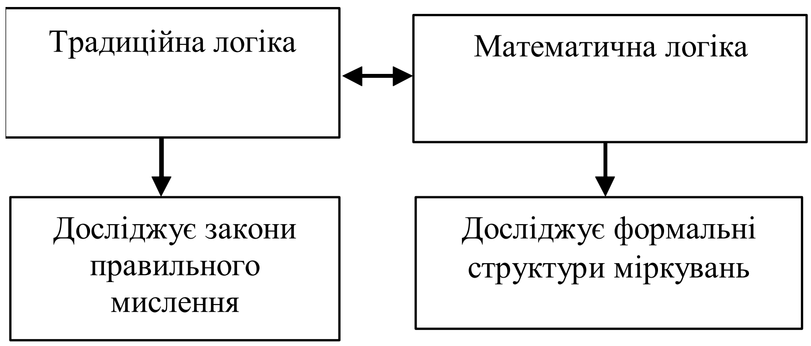


Рис.2 1. Структура логіки.

Логіка, вивчаючи правильні міркування, оперує поняттями істинності і хибності. Правильне міркування або висловлювання покладається істинним, неправильне – хибним.

Хоча в звичайному житті ми користуємося великим числом відтінків (модальностей), для формального апарату представлення знань нам досить двозначної логіки.

Методологічні основи формальної логіки полягають у виконанні декількох принципів.

Принцип тотожності. Істинність фактів, що лежать в основі висловлювань і міркувань, встановлюється на підставі відомих законів, спостережень. Якщо істинність будь-якого факту встановлено, то вона не піддається сумніву і не змінюється в процесі міркування. Це означає також, що один і той же термін використовується завжди в одному і тому ж сенсі.

Принцип несуперечності означає, що, стверджуючи що-небудь, не можна заперечувати те ж саме. Один і той же факт (вислів) не може бути одночасно істинним і хибним.

Наприклад, висловлювання Сократа *«Я знаю, що я нічого не знаю»* суперечливо, оскільки одночасно стверджує і спростовує один і той же факт: якщо Сократ знає, що він нічого не знає, то він не знає також і цього. Відповідно до принципу несуперечності, з розгляду виключаються такі висловлювання, істинність або хибність яких не може бути встановлена, наприклад, висловлювання про цирульника, який повинен (або не повинен) голити самого себе, висловлювання крітянина про те, що все крітяни – брехуни, і інші семантичні парадокси.

Принцип виключення третього. Не можна одночасно відкидати висловлювання і його заперечення. Будь-яке висловлювання може бути або істинним, або хибним, – третього не дано.

Принцип достатньої підстави. Будь-яке висловлювання має бути обгрунтовано, тобто істинність твердження не можна приймати на віру. Якщо твердження виводиться з будь-яких суджень, даних, фактів – *підстав,* то їх повинно бути досить для встановлення істинності твердження.

Можна ще вказати на одну дуже важливу властивість - **монотонностіь достовірних міркувань.** Якщо істинність деякого висловлювання *(A)* вже встановлена, то додавання нових фактів *(B)* не змінює істинності *A*.

2.2. Основні поняття

Судження, або висловлювання – це думка, в якій стверджується наявність або відсутність якихось фактів або зв’язків між фактами. Висловлення виражаються розповідними реченнями.

Визначення 2.1. *Просте висловлювання* – це просте розповідне речення, відносно якого можна однозначно сказати, істинне воно чи хибне.

Питальні або окличні речення висловлюваннями не є.

Логічні значення *Істинно* $(True)$ чи *Хибно* $(False)$ будемо позначати відповідно T и F.

Приклади.

«Київ – столиця України» – істинне висловлювання, воно має значення $True$ («істинно»). «5> 10» – хибне висловлювання, воно має значення $False$ («хибно»). «Все люди смертні» – істинне висловлювання. «Деякі люди– юристи» – істинне висловлювання.

Кожне просте висловлювання позначають символами латинського алфавіту (з індексами або без індексів), які називають *пропозиціональними символами:* $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{A}_1, \mathrm{A}_2 \ldots$

Складні висловлювання складаються з простих за допомогою сполучників «ні», «і», «або», «якщо…, то…», «тоді і тільки тоді». Цим сполучникам відповідають логічні операції: унарна операція заперечення $\neg$ («ні»), бінарні операції кон’юнкції & («і»), диз’юнкція $\vee$ («або»), імплікація $\rightarrow$ («якщо…, то…»), еквівалентність $\equiv$ («тоді і тільки тоді»). Символи операцій називаються *пропозиціональними зв’язками.*

Приклади:

«Лондон – столиця Англії» $\Leftrightarrow A$. «Лондон не є столицею Англії» $\Leftrightarrow \neg A$. «Найбільше місто Англії, Лондон (А), є її столицею (В)» $\Leftrightarrow A\&B$. «Населення Канади говорить англійською (А) або французькою мовою (В)» $\Leftrightarrow A \vee B$. «Якщо горобець – птаха (А), то у неї є крила (В)» $\Leftrightarrow A \rightarrow B$. «Тварина є птахою (А) тоді і тільки тоді, коли у неї є крила (В)» $\Leftrightarrow A \equiv B$.

Істинність або хибність складних висловлювань залежить від істинності або хибності простих висловлювань, які входять до складних, а також тим способом, якими вони комбінуються, тобто зв'язкою, яку використовують для побудови складного висловлювання. Кожна логічна зв'язка визначається своєю *таблицею істинності* (табл. 2.1, 2.2).

Таблиця 2.1

A	$\neg A$
F	T
T	F

Таблиця 2.2

A	B	$A \& B$	$A \vee B$	$A \to B$	$A \equiv B$
F	F	F	F	T	T
F	T	F	T	T	F
T	F	F	T	F	F
T	T	T	T	T	T

За допомогою зв'язки **заперечення** $\neg$ із стверджувальних утворюються заперечні висловлювання. Наприклад, якщо висловлювання А – «горобець – птаха» істинно, то висловлювання $\neg A$ – «горобець *не* птаха» - хибно, а висловлювання $\neg\neg A$ – «*Неправильно*, що горобець *не* птаха» еквівалентно висловлювання «горобець – птаха».

Кон'юнкція $A\&B$, що відповідає сполучникам «і», «а», істинна в тому і тільки в тому випадку, якщо істинні обоє з висловлювань, що входять в неї. Наприклад, висловлювання «на вулиці йде дощ (A) із сильним вітром (B)» виражається формулою $A\&B$. Кон'юнкція комутативна, тому висловлювання «на вулиці сильний вітер (B) і дощ (A)», яке можна висловити формулою $B\&A$, еквівалентне першому. Проте, в звичайній мові подібні висловлювання не завжди еквівалентні, наприклад, висловлювання «Джейн вийшла заміж (A) і народила дитину (B)» і «Джейн народила дитину (B) і вийшла заміж (A)», еквівалентні в силу комутативності кон'юнкції, описують, очевидно, зовсім різні ситуації в житті. В формальній логіці з цим доводиться миритися.

За допомогою кон'юнкції формалізується принцип несуперечливості: $A\&\neg A \equiv F$ – **закон протиріччя.**

Диз'юнкція $A \vee B$, що відповідає сполучнику «або», є істинною в будь-якому випадку, коли істинне хоча б одне висловлювання, що входить в неї, і хибна тільки у випадку, коли обидва простих висловлювання хибні. Наприклад, висловлювання «двічі два – чотири (A) або п'ять (B)» виражається формулою $A \vee B$ та є істинним. Диз'юнкція формалізує **закон виключення третього**: $A \vee \neg A \equiv T$.

Імплікація $A \to B$ виражає логічний (часто – причинно-наслідковий) зв'язок між висловлюваннями A і B та формалізує звичайний умовивід, в якому з посилки (антецедента) A випливає висновок (консеквент) B. Таблиця істинності імплікації відображає правильні (достовірні) умовиводи.

Формальне висловлювання **принципу тотожності** можна записати так: $\boldsymbol{A \to A}$ *(A є A)*. Формальне висловлювання принципу монотонності достовірних міркувань: $\boldsymbol{A \to (B \to A)}$.

Зазвичай імплікація $A \to B$ висловлює деяке правило, яке означає необхідність істинності посилки A для того, щоб виконувалась істинність висновку B, наприклад: «якщо воду нагріти до 100 градусів (A), то вона закипить (B)»; «щоб отримати диплом (B), потрібно закінчити інститут (A)»; «якщо кішка перебіжить дорогу (A), то трапиться неприємність (B)»; «якщо на небі хмари (A), може піти дощ (B)».

Еквівалентність $A \equiv B$ стверджує рівнозначність (рівносильність, тотожність) двох тверджень A і B; вона істинна тоді і тільки тоді, коли істинностні значення A і B співпадають.

Еквівалентність $A \equiv B$ стверджує не тільки необхідність умови A для того, щоб було істинно B, але і достатність цієї умови, тобто $A \equiv B = (A \to B)\ \&\ (B \to A)$, где $A \to B$ виражає *необхідність* A, а $B \to A$ – *достатність* A. Наприклад, стверджения: «птахи літають над морем – земля близько», – виражає одночасно два стверджения: необхідність умови – «якщо птахи літають над морем, то близько земля», і достатність: «якщо земля близько, то птахи літають над морем».

За допомогою логічних зв'язок складні висловлювання можна записати у вигляді формули, яку називають *пропозиціональною формулою.*

Визначення 2.2.

- Кожна пропозиціональна буква є формула.
- Якщо A і B – формули, то формулами є: $(\neg A)$, $(A \ \& \ B)$, $(A \vee B)$.
- Інших формул немає.

При записі формул прийнято наступне: зовнішні дужки можна опускати; встановлено пріоритет операцій: $\neg$, &, $\vee$, $\rightarrow$, $\equiv$. Логічні зв'язки для імплікації $\rightarrow$ та еквівалентності $\equiv$ вводяться для скорочення запису формул:

$A \rightarrow B = \neg A \vee B = \neg(A \ \& \ \neg B)$, $A \equiv B = (A \rightarrow B) \ \& \ (B \rightarrow A)$.

Побудувавши формулу алгебри висловлювань, ми відволікаємося від її змістовного сенсу і оперуємо тільки поняттями істинності та хибності.

Приписування пропозиціональним літерам їх істинностних значень називається *інтерпретацією формули*. Множина всіх інтерпретацій формули утворює її таблицю істинності. Якщо виконати відображення $0 \Leftrightarrow F$, $1 \Leftrightarrow T$, то кожній пропозиціональній зв'язці буде відповідати булева операція, а кожній формулі алгебри висловлювань – булева формула, з чого випливає, що алгебра висловлювань є інтерпретацією булевої алгебри. В зв'язку з цим в ній зберігаються всі аксіоми і теореми булевої алгебри, зокрема й зображуваність формул алгебри висловлювань у вигляді ДДНФ і ДКНФ.

Визначення 2.3. Тотожно істинна формула називається *тавтологією*. Тотожно хибна формула називається *протиріччям*. Формула, яка приймає істинне значення хоча б на одній своїй інтерпретації, називається *здійсненною.* Формула, яка приймає на одних наборах істинні значення, а на якихось – хибні, називається *нейтральною.* Таким чином, формули розпадаються на три класи.

Дві формули називаються еквівалентними, якщо їх таблиці істинності співпадають.

Запис $\models A$ означає: «формула A – тавтологія».

Тавтології є *виділенними* формулами логіки висловлювань; саме вони представляють найбільший інтерес. Тавтологія істинна при будь-яких значеннях простих висловлювань, що входять в неї; таблиця істинності тавтології в кожному рядку містить значення «істинне». В звичайному житті ми також використовуємо поняття тавтології: це твердження, яке не несе ніякої інформації, оскільки посилка і його висновок стверджують одне й те саме, наприклад: «Сніг білий, тому, що він білий», «Масло – масляне». Такі висловлювання виражаються формулою: $A \to A$, тотожню істинність якої не складно перевірити: якщо $|A| = F$, то $F \to F = T$, якщо $|A| = T$, то $T \to T = T$. Твердження: «Хворий або помре, або виживе», сакраментальна фраза «Бути чи не бути» – також тавтологія, так як вони можуть бути виражені тотожно істинною формулою: $A \vee \neg A$.

В формальній логіці поняття тавтології включає в себе всі тотожно істинні твердження. Вони формалізують правильні схеми міркувань. Наприклад, закони де Моргана – це тавтології: $\models \neg(A \vee B) \equiv (\neg A \ \& \ \neg B)$ (наприклад, *«неправильно, що цей злочин скоїли А і В»* еквівалентно висловлюванню: *«ні А, ні В не скоювали цього злочину»*); $\models \neg(A \ \& \ B) \equiv (\neg A \vee \neg B)$ (наприклад, *«неправильно, що А і В разом приймали участь в пограбуванні»* еквівалентно *«або А, або В, або обоє вони в пограбуванні не брали участь»*).

Проблема можливості розв'язання в алгебрі висловлювань полягає в тому, щоб відшукати ефективну процедуру (алгоритм), за допомогою якої для кожної формули логіки висловлювань можна встановити, чи є вона тавтологією, чи ні.

Очевидно, що така процедура для формул логіки висловлювань існує: це побудова таблиць істинності.

Приклади.

1. Як приклад розглянемо закон ствердження консеквента: $\models A \to (B \to A)$. (***принцип монотонності достовірних міркувань***). Оскільки кожен рядок таблиці істинності (табл. 2.3) цієї формули містить значення Т, формула є тавтоогією. Також, $A \to (B \to A) = \neg A \vee \neg B \vee A = T \vee \neg B = T.$

Таблиця 2.3

A	B	$B \to A$	$A \to (B \to A)$
F	F	T	T
F	T	F	T
T	F	T	T
T	T	T	T

2. **Метод редукції (зведення до протиріччя)** – це спосіб скорочення переборів при складанні таблиці істинності. В якості приклада доведемо, що формула $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ – тавтологія.

Припустимо, що це не так, тобто існує така інтерпретація, на якій формула приймає хибне значення:

$$|\overbrace{(A \to (B \to C))}^{T} \to \overbrace{(\underbrace{(A \to B)}_{T} \to \underbrace{(A \to C)}_{F})}^{F}| = F.$$

Маємо систему рівнянь :

1) $|A \to (B \to C)| = T,$

2) $|A \to B| = T$,

3) $|A \to C| = F$.

З 3) випливає: $|A| = T$, $|C| = F$. Підставимо ці значення в 2): $|T \to B| = T$, звідки $|B| = T$. Підставимо знайдені значення $|A| = T$, $|C| = F$, $|B| = T$ в 1): $|T \to (T \to F)| = |T \to F| = F$. Отримане значення суперечить умові 1), а отже, не існує такої інтерпретації, на якій формула приймає хибне значення, тобто вона є тавтологією.

3. Перевіримо, чи є така формула тавтологією: $(A \to (B \to C)) \to (A \vee B \to C)$.

Припустимо, що $|(A \to (B \to C)) \to (A \vee B \to C)| = F$. Тоді

1) $|A \to (B \to C)| = T$,

2) $|A \vee B| = T$,

3) $|C| = F$.

Із 2) випливає: а) $|A| = T$, $|B| = T$; б) $|A| = T$, $|B| = F$; с) $|A| = F$, $|B| = T$. Підставляючи значення $|A| = T$, $|B| = T$ в 1), отримаємо протиріччя: $|T \to (T \to F)| = F$. Але це ще не доводить, що формула є тавтологією. Розглянемо інші значення: б) $|A| = T$, $|B| = F$. Підставляючи ці значення в 1), маємо: $|T \to (F \to F)| = T$. Таким чином, при інтерпретації $|A| = T$, $|B| = F$, $|C| = F$ формула приймає хибне значення, а отже, вона не є тавтологією.

2.3. Логічний наслідок

Визначення 2.4. Якщо A і B – формули, то кажуть, що *B логічно випливає з A*, чи *A логічно призводить B*, якщо на всіх інтерпретаціях, де A приймає істинне значення, B також приймає істинне значення. Це позначають як $A \models B$ чи $A \Rightarrow B$.

Теорема 2.1. Логічний наслідок A |= B виконан тоді і тільки тоді, коли формула $A \to B$ – тавтологія.

Доведення. Нехай логічний наслідок A |= B виконан. Це означає, що на всіх інтерпретаціях, на яких формула |A| = T, формула |B| = T, тоді, |A→B| =T.

Якщо формула |A| = F, то | A $\to$ B | = T незалежно від значення B, то $A \to B$ – тавтологія.

Припустимо, що формула $A \to B$ – тавтологія. Тоді не існує такої інтерпретації, на якій | A $\to$ B | = F. Отже, якщо формула | A | = T, то і | B | = T, що відповідає визначенню логічного наслідку, тобто A | = B. ◊

Визначення 2.5. Формула B логічно випливає з формул *A1, A2, ..., An*, якщо на всіх тих інтерпретаціях, на *яких A1, ..., An* приймають істинні значення одночасно, формула *B* також приймає істинне значення. Це позначається так: *A1, ... An* |= *B.*

Теорема 2.2. *A1, A2, ..., An* |= *B* тоді і тільки тоді, коли $\models A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_n \to B$.

Довести самостійно.

Визначення 2.5. A |= B і B |= A, то формула A логічно еквівалентна формулі B. Це позначається як $A \Leftrightarrow B$, чи $A \equiv B$.

Якщо формула A логічно еквівалентна формулі B, то $A \equiv B$ – тавтологія.

Приклад. Перевіримо логічний наслідок: $A \to B,\ A \to \neg B \models \neg A$ (правило доведення до абсурду) і тавтологію: $\models (A \to B) \,\&\, (A \to \neg B) \to \neg A$. Позначимо тавтологію через *E.* Побудуємо таблицю істинності (табл. 2.4).

Таблиця 2.4

A	B	$A \to B$	$A \to \neg B$	$(A \to B) \,\&\, (A \to \neg B)$	$\neg A$	E
F	*F*	*T*	*T*	*T*	*T*	*T*
F	*T*	*T*	*T*	*T*	*T*	*T*
T	*F*	*F*	*T*	*F*	*F*	*T*
T	*T*	*T*	*F*	*F*	*F*	*T*

Як бачимо, на тих наборах, на яких посилки $A \to B$, $A \to \neg B$ приймають істинне значення одночасно, формула ¬A також приймає істинне значення. Отже, логічний наслідок виконаний. З іншого боку, формула

$(A \to B)$ & $(A \to \neg B) \to \neg A$ є тавтологією, що також є доказом дійсності логічного наслідку.

Логічні наслідки встановлюють *правила виведення* істинних висновків з істинних посилок. Кажуть, що логічне слідування (правило виведення) *зберігає істинність.* Правила виведення, що зберігають істинність, називаються *достовірними (дедуктивним) міркуваннями.*

2.4. Метатеореми про тавтології

Теорема 2.3 (*правило modus ponens*). Якщо A – тавтологія і $A \to B$ – тавтологія, то B – тавтологія, тобто якщо $|= A$ *і* $|= A \to B$, то $|= B$.

Доведення. Припустимо, що на деякій інтерпретації $| B | = F$. Тоді $| A \to B | = | A \to F | = T$ на тій самій інтерпретації (за умовою теореми). Отже, $| A | = F$, що неможливо, так як A - тавтологія. ◊

Правило *modus ponens* (скорочено MP) встановлює логічне слідування A, $A \to B \models B$ і називається ще правилом *відділення.*

Правило MP висловлює елементарний акт дедукції. Імплікації $A \to B$, яка за визначенням має сенс «якщо A, то B», можна інтерпретувати як правило, в якому A є «причиною», а B - «наслідком». Тоді правило MP говорить про те, що наслідок B настає при виконанні умови A, тобто при істинності посилки.

Наприклад, формула $A \to B$ може висловлювати таке правило: «*якщо на світлофорі горить зелене світло, то можна переходити дорогу*». Ми чекаємо на зелене світло на світлофорі, щоб перейти дорогу, тобто ми неявно користуємося правилом MP: коли посилка стає справжньою (зелене світло), то істинно і висновок (можна переходити дорогу). Тим самим ми виконуємо елементарний акт дедукції: з істинності посилки ми одержуємо дійсний висновок.

Теорема 2.4 (*правило підстановки*). Якщо A – тавтологія, яка містить пропозіційні змінні $a_1, a_2, ..., a_n$, то формула B, отримана з A заміною кожного входження a_i на деяку формулу A_i, також буде тавтологією.

Доведення. Нехай задано істиннісний розподіл для пропозіційних літер, що входять в В. Формули A_i для цього розподілу приймуть деякі значення $\delta_1, \delta_2, ..., \delta_n$, де δ_i є T або F. Якщо такий же розподіл надати формулами $A_1, ... , A_n$, то значення формули B співпаде зі значенням формули A. Але так як A – тавтологія, то значення буде T при будь-якому істинносному розподілі літер, що входять в B. Отже, B – тавтологія. ◊

Приклад. Формула $A \to (B \to A)$ - тавтологія. Замінимо A на $A \vee B$. Отримаємо нову тавтологію: $| = A \vee B \to (B \to A \vee B)$. Таким чином, кожну тавтологію можна розглядати як схему, з якої за допомогою підстановки можна отримати безліч тавтологій.

Теорема 2.5 (*правило еквівалентної заміни*). Якщо B можна отримати з A підстановкою формули B_1 замість одного або декількох входжень підформули A_1 в A, то $((A_1 \equiv B_1) \to (A \equiv B))$ є тавтологія, і, отже, якщо A_1 *і* B_1 логічно еквівалентні, то A і B також логічно еквівалентні.

Іншими словами, якщо є тавтологія A, і в ній є підформула A_1, і, якщо замінити A_1 на еквівалентну їй формулу B_1, то отримана формула B буде еквівалентна A.

Доведення. Розглянемо довільний істинносний розподіл змінних, що входять в A, B, A_1, B_1. Якщо при цьому розподілі A_1 *і* B_1 мають різні значення, то $| A_1 \equiv B_1 | = F$ і, отже, $((A_1 \equiv B_1) \to (A \equiv B))$ дорівнює T. Якщо ж A_1 *і* B_1 мають однакові значення, то однакові істинносні значення будуть у A *і* B, так як B відрізняється від A тільки тим, що деякі входження підформули A_1 замінені в ній на B_1, яка має те ж саме істинносне значення. Отже, в цьому випадку, якщо $| A_1 \equiv B_1 | = T$, то і $| A \equiv B | = T$, і $((A_1 \equiv B_1) \to (A \equiv B))$ є тавтологія. ◊

Приклад. У тавтології $A \to (B \to A)$ замінимо підформулу $B \to A$ на еквівалентну їй формулу $\neg B \vee A$, отримаємо нову тавтологію $A \to \neg B \vee A$. Тавтологією також буде формула: $A \to \neg (B \& \neg A)$, так як $B \to A$ еквівалентна $\neg (B \& \neg A)$. Розглянуті метатеореми дають можливість отримувати нові тавтології і нові правила виведення з уже наявних.

2.5. Формалізація і рішення логічних задач

Мова алгебри висловлювань використовується для формалізації речень природної мови і доказів логічних слідувань. Для цього кожне просте висловлювання позначається пропозіціональною літерою; складне висловлювання записується як формула, в якій зв'язки з природної мови замінюються пропозіціональними зв'язками (як зазначено в 2.2).

Двозначна класична логіка корисна в повсякденному житті, навіть, якщо доводиться визнати, що в припущеннях, наприклад, про двозначність, є елемент свавілля. Для вирішення будь-якої інтелектуальної проблеми не слід нехтувати логікою. Наприклад, якщо розглядається якась ймовірнісна задача, то перш за все можна знайти наслідки, що випливають з набору гіпотез, без розгляду їх ймовірнісного (змістовного) сенсу. Посилання на модальну або інтуїционістську логіку дають зрозуміти, що в деяких ситуаціях потрібна інша, не класична логіка

Зв'язкою «еквівалентність» замінюються речення природної мови наступної природи: *якщо - і - тільки - якщо, тоді і тільки тоді, рівносильно.* «Імплікація» замінює: *якщо - то, в разі А має місце В, А – тільки, якщо В, В - якщо А, А тягне В, для В досить А, для А необхідно В*. Заміна на «кон'юнкцію» - *А і В , не тільки А - а й В, В - хоча і А, В - незважаючи на А, як А - так і В, А разом з В, А - в той же час як В.*

Переклад ***А & В*** найбільш простий, якщо в ньому не бере участь причина часу (про що йшлося вище). Багато смислових відтінків (незважаючи, хоча, але і) зникають при формалізації.

«*Диз'юнкція*» - А чи В, А - якщо не В, А чи В або обидва. Труднощі перекладу «*диз'юнкції*» полягають в двозначності певних термінів. Якщо в меню ресторану вказано «*чай або кава - безкоштовно*», то ми не здивуємося, якщо, зажадавши, і те й інше, отримаємо збільшений рахунок. У той же час, коли оголошення говорить, що «*пожертвування приймаються в церкві або школі*», то ми не думаємо, що принесене в церкву відкинуть через те, що ми вже пожертвували в школі.

Якщо хочуть висловити виняткове «*або»*, то повинні додати до «*або*», наприклад, фразу «*але не те й інше разом*» - формалізується $\neg$ (A & B) & ($A \vee B$).

Виразам «*A, якщо не B*», «*A, крім випадку, коли B*» надається сенс такого, що наявність *B* звільняє від відповідальності стверджувати істинність *A*. Тобто, якщо не $\neg B$, то *A*, а якщо *B*, то про *A* нічого не говориться ($\neg B \to A$ = $\neg A \to B$).

У звичайній мові не вживаються дужки для вказівки сполучуваності частин складного речення. «*Якщо Джонс присутній або якщо Вільямс висловиться за нашу пропозицію і Старк не стане заперечувати, то наша пропозиція буде прийнята*» («стратити не можна помилувати» - на листі необхідно розставляти потрібні акценти знаками пунктуації). Формалізація цього речення наступна:

$\boldsymbol{D \vee W \,\&\, \neg S \to P}$ – формулі відсутні дужки, що не дозволяє однозначно сприймати інформацію відповідно з вихідним текстом (слід зазначити, що і правила побудови формул теж порушуються).

Приклад 2.1. Розглянемо логічне слідування: *Якщо не буде дощу, то ми поїдемо на пікнік. Якщо ми поїдемо на пікнік, то ми добре проведемо час. Дощу немає. Отже, ми добре проведемо час.*

Позначимо пропозіціональнимі буквами прості висловлювання: *P* - «*не буде дощу*»; *S* - *«поїдемо на пікнік»; R* - «*добре проведемо час*». Необхідно довести логічне слідування : $P \to S, S \to R, P \mid = R$.

Доведення.

1 спосіб. Довести тавтологію: $\mid = (P \to S) \,\&\, (S \to R) \,\&\, P \to R$. Для цього досить побудувати таблицю істинності.

2 спосіб. Доведення від протилежного (метод редукції).

Припустимо, що існує така інтерпретація, на якій всі посилки приймають істинне значення, а наслідок - хибне, тобто $\mid P \to S \mid = T$, $\mid S \to R \mid = T$, $\mid P \mid = T$. Припускаємо $\mid R \mid = F$. Тоді з $\mid S \to R \mid = \mid S \to F \mid = T$ випливає, що $\mid S \mid = F$;

з $|P \to S| = |P \to F| = T$ випливає, *що* $|P| = F$, що суперечить третьої посилці | P | = T. Це протиріччя доводить логічне слідування.

3 спосіб. Побудова формального логічного висновку.

Побудуємо логічний висновок, використовуючи відомі правила виведення. Логічний висновок записується зазвичай як пронумерована послідовність формул, праворуч від кожної формули записується коментар, який вказує, на якій підставі формула включена в цю послідовність. Посилки виведення зазвичай позначаються літерою Г (гіпотеза).

1. $P \to S$ Г1
2. $S \to R$ Г2
3. P Г3
4. S правило МР (3,1)
5. R правило МР (4,2)

Остання формула в цьому висновку є логічним наслідком посилок *Г1, Г2, Г3*.

В силу того, що правило МР зберігає істинність, кожна формула, яка бере участь у виведенні, істинна при істинності посилок Г1, Г2, Г3. Тому, якщо з'єднати символом $\to$ формули, так, щоб посилка імплікації передувала висновку в цьому виведенні, то отримана формула також буде істинною, і, отже, вона є логічним наслідком посилок.

Тому з даного висновку ми можемо отримати логічне виведення:

$P \to S, S \to R \mid = P \to R.$

Це логічне слідування відповідає правилу виведення, яке називають правилом силогізму:

$A \to B, B \to C \mid = A \to C.$

Приклад 2.2. *Через снігопад (А) машина потрапить в затор (В):* $A \to B$. *У затор машина не потрапила* ($\neg B$). *Отже, снігопаду не було* ($\neg A$). *Доведемо* $A \to B, \neg B \mid = \neg A$.

Припустимо, що $| A \to B | = T$, $| \neg B | = T$ і $| \neg A | = F$, тоді $| A \to B | = | T \to F | = F$. Отримане протиріччя доводить логічне слідування.

Це логічне слідування відповідає правилу виведення ***modus tollens (MT):***

$A \to B, \neg B | = \neg A.$

Приклад 2.3. *Решітка є модулярною (В) в тому випадку, якщо вона дистрибутивною (А):* $A \to B$. *Але ми знаємо, що якщо решітка немодулярна* $(\neg B)$*, то вона і недистрибутивна* $(\neg A)$*:* $\neg B \to \neg A$.

Припустимо, що $| A \to B | = T$, $| \neg B \to \neg A | = F$, тоді $| \neg B | = T$, $| \neg A | = F$ і $| A \to B | = | T \to F | = F$. Отримане протиріччя доводить логічне слідування.

Це логічне слідування відповідає правилу виведення, яке називають правилом контрапозиції:

$A \to B | = \neg B \to \neg A$

Приклад 2.4. *Якщо долар росте (В), то ціни не зменшуються* $(\neg A)$. *Якщо долар падає* $(\neg B)$*, то ціни теж не зменшуються* $(\neg A)$*:* $B \to \neg A$ *і* $\neg B \to \neg A$. *Тобто, ціни не зменшуватися в будь-якому випадку.* Згідно із законом контрапозиції ми можемо отримати з умов задачі: $A \to \neg B$ *і* $A \to B$.

$\boldsymbol{A \to \neg B, A \to B | = \neg A.}$

Цей закон відображає правило приведення до абсурду. Якщо з посилки слідує одночасно істино і хибно, то сама посилка – хибна.

Приклад 2.5. Я не зможу заснути (A) тільки в тому випадку, якщо вип'ю каву (B). Як формалізувати це речення? Здається, що «якщо» знаходиться при висловлюванні B, то має бути $\boldsymbol{B \to A}$. Але сенс фрази полягає в наступному: «Якщо я не заснув, то з цього випливає тільки одне - я випив каву». Тобто, слід проводити формалізацію таким чином: $\boldsymbol{A \to B}$. Вживання «тільки», «крім» тощо змінює, напрямок імплікації.

Приведемо ще раз правила формального висновку:

$\boldsymbol{A \to B, A \models B}$	***modus ponens***
$\boldsymbol{A \to B, B \to C \models A \to C}$	***правило силогізму***

$\boldsymbol{A \to B \models \neg B \to \neg A}$	***правило контрапозиції***
$\boldsymbol{A \to B, \neg B \models \neg A}$	***modus tollens***
$\boldsymbol{A \to B, A \to \neg B \models \neg A}$	***правило приведення до абсурду***

Приклад 2.6. Отримати можливі логічні наслідки з даної множини посилок: *Якщо в городі - бузина (А), то в Києві - дядько (В). Коні їдять овес (С) або в городі - бузина. Якщо коні їдять овес, то Волга впадає в Каспійське море (Е). Волга не впадає в Каспійське море.*

Побудуємо дедуктивний формальний висновок:

1. $A \to B$	Г1
2. $C \vee A$	Г2
3. $C \to E$	Г3
4. $\neg E$	Г4
5. $\neg E \to \neg C$	правило контрапозиції (3)

(*Якщо Волга не впадає в Каспійське море, то коні не їдять овес)*

6. $\neg C \to A$	*еквівалентна заміна (2)*
7. $\neg E \to A$	правило силогізму (5, 6)

(Якщо Волга не впадає в Каспійське море, то в городі - бузина)

8. $\neg E \to B$	правило силогізму (1, 7)

(Якщо Волга не впадає в Каспійське море, то в Києві дядько)

9. B	МР (4,8)

(В Києві - дядько)

Таким чином, з даного набору посилок ми крок за кроком отримали всі можливі висновки.

Поняття про дедуктивний метод ми черпаємо, перш за все, з детективної літератури, зокрема, від Шерлока Холмса. Як відомо, Шерлок Холмс користувався при розкритті злочинів саме дедуктивним методом. Ось як він пояснює сутність дедуктивного методу:

«Мій принцип розслідування полягає в тому, щоб виключити всі явно неможливі припущення. Тоді те, що залишається, є істиною, якою б неправдоподібною вона не здавалася». Це міркування можна виразити такою схемою: $A \vee B, \neg A \models B$, що еквівалентно: $\neg A \to B, \neg A \models B$, тобто застосування правила MP.

В одному з оповідань про Шерлока Холмса склалася така ситуація: «Нам відомо, що злочинець не міг потрапити в кімнату ні через двері (A), ні через камін (B). Ми знаємо також, що він не міг сховатися в кімнаті (C), оскільки в ній ховатися ніде. Як же тоді він проник сюди? - Через дах (D)! - Без сумніву. Він міг проникнути в цю кімнату тільки через дах.»

Це міркування можна формалізувати так: $A \vee B \vee C \vee D, \neg A, \neg B, \neg C \models D$, що рівносильно: $\neg A \to (\neg B \to (\neg C \to D)), \neg A, \neg B, \neg C \models D$. Триразове застосування правила MP доводить цей логічний висновок.

Задачі (за теоремою 2.2.) відповідає формула:

$(A \vee B \vee C \vee D) \& \neg A \& \neg B \& \neg C \to D$, яка має бути тавтологією у випадку, коли висновок з цієї множини посилок є дійсним.

Не слід забувати, що логічний наслідок виконується тільки тоді, коли з істинних висловлювань слідує істинний висновок. Тому повинна існувати хоча б одна інтерпретація (тобто істиннісний розподіл літер, що входять в кожне висловлювання), на якій всі посилки істинні одночасно. Якщо такої інтерпретації не існує, то система посилок суперечлива і з неї виводиться будь-який висновок, тобто разом з деякою формулою A виводиться і її заперечення $\neg A$. З іншого боку, логічний наслідок може виявитися невиконаним, якщо посилок недостатньо для виведення потрібних висновків. У таких випадках, в залежності від змісту завдання, множину посилок може бути поповнено.

Приклад 2.7. *Четверо підсудних A, B, C і D. Встановлено наступне: якщо A винен, то B був співучасником; якщо B винен, то або C був співучасником, або A не винен; якщо D не винен, то A винен і C невинний; якщо B винен, то A винен. Хто з підсудних винен?*

Формалізуємо гіпотези: $A \to B$, $B \to C \vee \neg A$, $\neg D \to A \& \neg C$, $D \to A$. *Якщо перевірити множину посилок на несуперечливість, то отримаємо єдину можливість перетворення їх в істину – всі винні. Таку ж відповідь (A&B&C&D) буде отримано в результаті алгебраїчного спрощення формули відповідної задачі.*

2.6. Метод резолюцій

В основі методу резолюцій лежить процедура пошуку спростування, тобто замість доведення тавтології, що відповідає логічній задачі, доводиться те, що заперечення формули є протиріччям. Метод спростування для доказу логічного слідування полягає в наступному. Нехай виконується логічне слідування: $F1$, $F2$ |= G. Тоді |= $F1\&F2 \to G$ є тавтологією, і, отже, $|\neg(F1\&F2 \to G)| = |F1\&F2\&\neg G| \equiv F$. Оскільки за визначенням посилки $F1$, $F2$ істинні, формула $F1\&F2\&\neg G$ може обернутися в хибну тільки, якщо $|\neg G| = F$, тобто якщо $|G| = T$. Тоді логічне слідування виконано. В принципі процедура спростування формалізує метод редукції (так званий неформальний вивід).

Введемо декілька визначень.

Визначення 2.6. Диз'юнкція літер називається диз'юнктом, або clause (клаузою, клозом). Однолітерний диз'юнкт називається *одиничним диз'юнктом (фактом)*. Коли диз'юнкт не містить ніяких літер, його називають *пустим* диз'юнктом. Так як пустий диз'юнкт не містить літер, які могли б бути істинними за будь-яких інтерпретацій, то пустий диз'юнкт завжди хибний. Пустий диз'юнкт позначається символом □.

Правило резолюцій Робінсона. Якщо для будь-яких двох диз'юнктів C_1 і C_2 існує літера $L_1 \in C_1$ і контрарна їй літера $L_2 \in C_2$ ($L_2 = \neg L_1$), то викресливши L_1 з C_1 і L_2 з C_2 і побудувавши диз'юнкт з решти літер, отримаємо так звану *резольвенту* C_1 і C_2 : $C'_1 \vee C'_2$, де $C'_1 = C_1 \backslash L_1$, $C'_2 = C_2 \backslash L_2$.

Теорема 2.6. *Резольвента* C є логічним слідуванням C_1 і C_2, що містять контрарні літери L і $\neg L$: $L \vee C'_1, \neg L_1 \vee C'_2 \models C'_1 \vee C'_2$.

Доведення. Припустимо, що $|L \vee C'_1| = T$, $|\neg L \vee C'_2| = T$, $|C'_1 \vee C'_2| = F$. Тоді $|C'_1| = F$, $|C'_2| = F$. Якщо $|L \vee C'_1| = T$, то $|L| = T$, але $|\neg L \vee C'_2| = T$, отже, $|L| = F$. Отримане протиріччя доводить теорему.◊

Правило резолюцій є узагальненням багатьох відомих нам правил виводу. Наприклад, правило силогізму: $A \to B, B \to C \models A \to C$ може бути переписано у вигляді: $\neg A \vee B, \neg B \vee C \models \neg A \vee C$, що відповідає правилу резолюцій. Правило MP: $A, A \to B \models B$ може бути переписано у вигляді: $A \neg A \vee B \models B$, що також відповідає правилу резолюцій. Нарешті, закон протиріччя $A \& \neg A \equiv F$ рівнозначний правилу: $A, \neg A \models \square$, згідно якому, резольвента двох контрарних однолітерних диз'юнктів є пустий диз'юнкт.

Метод резолюцій стосовно вирішення задачі полягає в тому, щоб перевірити, чи містить множина диз'юнктів S, що складається з перетворених в КНФ посилок і заперечення висновка, порожній диз'юнкт $\square$. Якщо S містить $\square$, то множина S нездійсненна, якщо ні, то треба перевірити, чи може він бути отриманий з даної множини диз'юнктів. Іншими словами, необхідно знайти множину основних прикладів, які спростовують вихідну множину диз'юнктів. Ця процедура полягає в побудові резолютивного виведення і ґрунтується на правилі резолюцій Робінсона.

Визначення 2.7. *Резолютивний вивід* з множини диз'юнктів S є послідовність C_1, C_2, ..., C_k, така, що кожне C_i або належить S, або є резольвентою попередніх C_i. Якщо останній диз'юнкт $C_k = \square$, то множина диз'юнктів S є нездійсненною, а весь вивід називається спростуванням S. Якщо C_k не є пустим диз'юнктом і подальше застосування правила резолюцій неможливо, то множина S є здійсненною (але висновок в цьому разі не є наслідком з множини посилок, тобто, задача не розв’язується).

Приклад 2.8. Розглянемо задачу з прикладу 2.1. Необхідно перевірити логічне слідування в логіці висловлювань: $P \to S$, $S \to R$, $P \models R$. Складемо множину диз'юнктів S, для чого кожну формулу приведемо до КНФ, а від виснвка R візьмемо заперечення. Отримаємо:

1. $\neg P \vee S$
2. $\neg S \vee R$
3. P
4. $\neg R$
5. $\neg S$ резольвента 4, 2
6. $\neg P$ резольвента 5, 1
7. □ резольвента 3, 6

Приклад 2.9.

Які висновки можна зробити з наступних висловлювань:

" Повернувшись додому, Мегре подзвонив на набережну Орфевр:

- Є новини?

- Так, шеф. Надійшли повідомлення від інспекторів. Торранс встановив, що, якщо Франсуа був п'яний (F), то або Етьєн вбивця (E), або Франсуа бреше (L). Жуссе вважає, що або Етьєн вбивця, або Франсуа не був п'яний і вбивство сталося після опівночі (U). Інспектор Люка просив передати Вам, що якщо вбивство сталося після опівночі, то або Етьєн вбивця, або Франсуа бреше. Потім дзвонила ...

- Усе. Дякую. Цього достатньо. - Комісар поклав трубку. Він знав, що тверезий Франсуа ніколи не бреше. Тепер він знав все. "

Припустимо, що вбивця - Етьєн. Можна формалізувати задачу і застосувати метод резолюції для з'ясування істинності нашого припущення.

1. $F \rightarrow E \vee L$ *Умова задачі - список гіпотез*
2. $E \vee \neg F \& U$
3. $U \rightarrow E \vee L$
4. $\neg F \rightarrow \neg L$

Приведемо гіпотези в КНФ

1. $\neg E$ заперечення виводу
2. $\neg F \vee E \vee L$ гіпотеза 1
3. $E \vee \neg F$ гіпотеза 2 (видалення кон'юнкції)

4. $E \vee U$ гіпотеза 2 (видалення кон'юнкції)

5. $\neg U \vee E \vee L$ гіпотеза 3

6. $F \vee \neg L$ гіпотеза 4

7. $\neg F$ резольвента 1,3

8. U резольвента 1,4

9. $\neg U \vee L$ резольвента 1,5

10. $\neg L$ резольвента 6,7

11. $\neg U$ резольвента 9,10

12. $\square$ резольвента 8,11

Правило резолюцій – дуже потужний і зручний засіб логічного доведення.

Для побудови формального виводу необхідне знання деяких прийомів і теорем. Приведемо формальний вивід для вищевикладеної задачі:

1. $F \to E \vee L$ гіпотеза 1

2. $E \vee \neg F \& U$ гіпотеза 2

3. $U \to E \vee L$ гіпотеза 3

4. $\neg F \to \neg L$ гіпотеза 4

5. $\neg E \to \neg F$ гіпотеза 2 (видалення кон'юнкції)

6. $\neg E \to U$ гіпотеза 2 видалення кон'юнкції)

7. $\neg E \to E \vee L$ силогізм 3, 8

8. $\neg E \to L$ з 7 еквів.перетв. і ідемпотентність

9. $L \to F$ контрапозиція гіпотези 4

10. $\neg E \to F$ силогізм 8, 11

11. E довед.до абсурда 5, 10

Ми бачимо, що гіпотеза 1 є зайвою, вона не бере участь ні в доказі логічного слідування методом резолюцій, ні в побудові формального виводу.

Якщо ми використаємо гіпотезу 1, наприклад, зробимо силогізм з контрапозицією гипотези 4 ($L\rightarrow F$ та $F \rightarrow E\vee L$), то отримаємо $L \rightarrow E\vee L = =\neg L \vee E\vee L=T\vee E=T$, що є даремним кроком.

Приклад 2.10.

Побудуємо резолютивний вивід для задачі з прикладу 2.6. Але трохи змінемо умову, щоб розглянути приклад отримання заперечення висновку для іншої форми, ніж в попередній задачі.

Якщо в городі - бузина (A), то в Києві - дядько (B). Коні їдять овес (C) або в городі - бузина. Якщо коні їдять овес, то Волга впадає в Каспійське море (E). ~~Волга не впадає в Каспійське море~~. *(останню посилку викреслюємо). Висновок для перевірки:* Якщо Волга не впадає в Каспійське море, то в Київі – дядько *(формальний вивід зупинився би на п.8 – див вище розв'язок прикладу 2.6).*

Формалізація:

1. $A \rightarrow B$	Г1	$(\neg A \vee B)$
2. $C \vee A$	Г2	так вже і є диз'юнкт
3. $C\rightarrow E$	Г3	$(\neg C\vee E)$
$\neg E\rightarrow B$	висновок	

Візьмемо заперечення висновку:

$\neg(\neg E\rightarrow B)= \neg(\neg\neg E \vee B)= \neg(E \vee B)= \neg E \& \neg B.$

Ми отримали два однолітерних негативних диз'юнкта: $\neg E$ *та* $\neg B$.

Запишемо множину диз'юнктів, яку будемо перевіряти на несуперечлівість: $\{\neg A \vee B, C \vee A, \neg C\vee E, \neg E, \neg B\}$.

1. $\neg E$ *заперечення висновку*
2. $\neg B$ *заперечення висновку*
3. $\neg A \vee B$ *Г1*
4. $C \vee A$ *Г*
5. $\neg C\vee E$ *Г3*
6. $\neg C$ *резольвента 1,5*
7. A *резольвента 4,6*

8. B *резольвента 3,7*

9. □ резольвента 2,8

Пустий диз’юнкт отримано, система суперечлива, висновок – дійсний.

Питання для самоконтролю до розділу 2

1. Назвати принципи формальної логіки та записати їх формальне визначення.
2. Просте та складне висловлювання. Зміст пропозиціональних зв’язок.
3. Що означає – ввести логічні операції?
4. Метатеореми о тавтологіях.
5. Основні логічні закони.
6. Основна ідея методу резолюції.
7. Пустий диз’юнкт та його значення.
8. Правило резолюцій Робінсона та визначення резольвенти.
9. Теорема о логічном слідуванні.
10. Методи доказу істиності висновку з множини посилань.

3. ФОРМАЛЬНІ ТЕОРІЇ

Числення - система вивчення тих або інших областей об'єктивного світу, в якій предметам якої-небудь певної області ставляться у відповідність матеріальні знаки (цифри, букви і тому подібне), з якими потім чисто формально за прийнятими в системі точними логічними правилами виконуються операції, необхідні для вирішення поставленої мети.

Математика, що виникла 6 тисяч років тому в стародавньому Єгипті та Вавилоні будувалась, насамперед, як числення. Тільки у III ст. до н. е. Евклід уперше побудував математику у вигляді аксіоматичної теорії. Але і в сучасній школі вивчення математики розпочинається з нумерації і чотирьох дій арифметики, тобто з оперування зі знаками (цифрами).

Аксіоматична теорія – теорія, яка побудована з скінченного числа постулатів чи аксіом, з яких за допомогою заданих логічних правил виведення дедуктивно можуть бути отримані усі інші універсально загальнозначущі, тобто змістовно-істинні речення (теореми), сформульовані на мові цієї теорії.

Іноді кажуть, що кожна аксіоматична теорія будується на двох положеннях:

1) на множині початкових постулатів або аксіом і множині доказових висловлювань, тобто теорем, що виводяться логічним шляхом з аксіом;

2) на логіці, яка дає правила, по яких з аксіом виводяться теореми.

Звідси витікає, що потрібно знати не лише аксіоми аксіоматичної теорії, аналізом яких зазвичай займаються досить докладно, але і формальну логіку, як традиційну, так і математичну.

Первинні терміни, які входять у формулювання аксіом, приймаються в аксіоматичній теорії без визначення. Якщо теореми логічно правильно виводяться з аксіом, що є істинами, то і теореми є істинами, оскільки з істинних положень за умови дотримання правил логіки завжди виходять істинні висновки.

У формалізованих і змістовних аксіоматичних теоріях застосовуються спеціальні наукові мови. При побудові формалізованих аксіоматичних теорій логіки спираються на мову системи, яка включає алфавіт початкових символів, а також правила утворення правильно побудованих виразів (формул). За допомогою цієї мови зазвичай записуються аксіоми. Для формулювання правил виведення звертаються до так званої метамови (мова, на основі якої проходить дослідження іншої мови, що називається об'єктною; можна привести аналогію – метамовою є рідна для вивчення іноземної).

Операції з символами здійснюються за такими правилами, які припускають врахування лише формальних властивостей знакових виразів, що фігурують в системі.

Доведенням в аксіоматичній теорії вважається кінцева послідовність формул цієї теорії, кожна з яких є або аксіомою, або безпосередньо виводиться з попередньої послідовності формул за загальноприйнятими правилами логіки. Останнім висловлюванням в цій послідовності формул, що називається доведенням, і виступає теорема, яку і потрібно було довести.

3.1. Визначення формальної теорії

Теорія називається *змістовною,* якщо існує яка-небудь інтерпретація об'єктів, операцій, символів теорії. Алгебра висловлювань є змістовною теорією, оскільки кожен символ в ній інтерпретується як просте висловлювання, істинне або хибне, формули також можуть бути істинними або хибними, і, попри те, що ми відволікаємося від змістовного сенсу висловлювань, усі міркування ведуться в термінах істинності і хибності. На відміну від змістовних теорій, формальна теорія є множиною символів і відношень між ними, яким не приписується ніякого змістовного сенсу. Згадуючи про завдання логіки, можна сказати, що вирази і формули формальної теорії представляють "чисті" схеми міркувань, яким можна надавати найрізноманітніший сенс, тобто будувати різні моделі теорії.

Формальні теорії будуються за наступними правилами.

Задається зчисленна множина символів, що називається *алфавітом* теорії. З цієї множини символів будуються зчисленні послідовності, які називаються *словами* або *виразами* даної теорії. З множини виразів виділяють *правильно побудовані вирази – формули*. З множини формул виділяють підмножину *аксіом*. Між формулами теорії визначають відношення – *правила виведення* теорії. Правила виведення дозволяють з множини аксіом виводити *теореми*. Таким чином, формальна теорія є аксіоматичною теорією.

Визначення 3.1. *Доведенням* теореми називають послідовність формул $A_1,\ \dots,A_n$, кожна з яких або є аксіомою, або отримана з попередніх за правилами виведення цієї теорії. Остання формула називається *теоремою* цієї теорії. Іншими словами, *теорема* - це формула, що виводиться з множини аксіом за правилами виведення, визначеними в цій теорії. Запис $|\!-_J A_n$ означає, що формула A_n виводиться в теорії J, тобто є теоремою теорії J.

Визначення 3.2. *Виведенням* формули A_n з множини гіпотез Г називається послідовність формул $A_1,\ \dots,A_n$, кожна з яких є або аксіомою, або гіпотезою з множини Г, або отриманою з попередніх по *правилах виведення*. Формула $B = A_n$ називається такою, що виводиться з множини гіпотез Г. Означення $\Gamma \vdash_J A_n$ означає, що формула A_n виводиться з множини гіпотез Г в теорії J.

Доведення відрізняється від виведення тим, що в ньому не використовуються гіпотези, тому теорему можна визначити як формулу, що виводиться з порожньої множини гіпотез.

Властивості виведення:

1. Якщо $\Gamma \vdash B$ і $\Gamma \subset \Delta$, то $\Delta \vdash B$.

Ця властивість називається властивістю *монотонності* виведення. Вона означає, що формула B буде як і раніше виводитися, якщо до множини гіпотез Г, з яких виводиться формула B, додати інші гіпотези, тобто розширити множину гіпотез Г до Δ.

2. $\Gamma \vdash B$ тоді і тільки тоді, коли існує $\Delta \subseteq \Gamma$, таке, що $\Delta \vdash B$. Це властивість *повноти* множини гіпотез: для того, щоб формула B виводилася з множини гіпотез Г, необхідно і достатньо, щоб в Г існувала підмножина $\Delta \subseteq \Gamma$, з якої виводиться формула B. Іншими словами, не усі гіпотези із заданої множини гіпотез Г обов'язково повинні використовуватися в процесі виведення, – деякі можуть виявитися зайвими, проте, заданих гіпотез повинно бути досить для виведення B.

3. Якщо $\Delta \vdash A$, і для кожного $B_i \in \Delta$, $\Gamma \vdash B_i$ то $\Gamma \vdash A$. Це властивість *транзитивності* виведення. Воно дозволяє використати раніше доведені теореми (виведення), не повторюючи усього списку формул, що становлять доведення (виведення). Тому раніше доведені теореми і виведення можуть використовуватися в інших доведеннях як схеми, в яких кожне входження змінної може заміщатися довільною формулою.

Аксіоматична теорія (АТ) називається такою, що вирішує проблему можливості розв'язання, якщо в ній є ефективна процедура (алгоритм), що вирішує у рамках цієї теорії певний круг проблем, і такою, що не вирішує, якщо немає такого алгоритму. При цьому потрібно зважати на те, що АТ будується з таким розрахунком, щоб теореми, що доводяться, при логічній інтерпретації виявлялися істинними.

Змістовні і формальні аксіоматичні системи можуть задовольняти одночасно різним системам об'єктів, які є для них моделями.

АТ має бути несуперечливою, тобто щоб в ній не можна було вивести теорему і одночасно її заперечення. Теорія, в якій можна одночасно вивести А і $\neg A$ на підставі прийнятих в ній правил, не має ніякої цінності, оскільки вона не в змозі відобразити відмінності між істиною і хибністю. Слід зауважити, що питання про несуперечність формальної системи не можна вирішити засобами, які формалізуються в цій же системі. Як показав Гедель, тут потрібно звертатися до сильніших логічних засобів.

Іншою вимогою до АТ є вимога незалежності системи її аксіом. Аксіома вважається незалежною, якщо її не можна вивести з інших аксіом, і якщо її виключення з системи аксіом зменшує запас теорем. Система аксіом називається незалежною, якщо кожна з аксіом, що входять в неї, є незалежною. Недоліком залежної системи є наявність "зайвих" аксіом, які не посилюють її ефективність.

Можна привести загальноприйнятий метод доведення незалежності системи аксіом: для доведення того, що аксіома A є незалежною по відношенню до деякої системи аксіом S, до системи S приєднують $\neg A$ і намагаються знайти об'єкти, що задовольняють S і $\neg A$. Якщо це можна зробити, то система, що складається з S і $\neg A$, несуперечлива, тоді A є незалежною по відношенню до системи S, і система S може бути поповнена аксіомою A .

АТ є повною, тоді і тільки тоді, коли усі теореми можуть бути отримані з цієї системи аксіом. Система аксіом є такою, що не можна поповнити, якщо приєднання до неї якої-небудь формули призводить до протиріччя. Правда неможливість поповнення (як і повнота) теорії не вважається такою ж обов'язковою якістю теорії, як несуперечність.

Логічне числення будується за допомогою формального апарату. Окремі напрями логічних досліджень розвиваються незалежно від потреб науки. Практика дає багато прикладів того, як той або інший логік бере деякі початкові припущення, формальним чином будує логічні числення і встановлює наслідки, які можна вивести з прийнятих початкових припущень. Створивши таким чином систему логічного числення, дослідник починає відшукувати відповідні інтерпретації і можливі застосування її до змістовних теорій. Так сталося з алгеброю Буля – спочатку з'явилося числення висловлювань, а потім воно було інтерпретоване на релейно-контактні схеми. Це відноситься до цілого ряду багатозначних логік, які побудовані чисто формальним чином, і які ще не інтерпретовані, але багато з яких знайдуть свої «застосування».

У математичній логіці є декілька взаємопов'язаних числень - числення висловлювань, числення класів, числення предикатів, числення відношень. Розробкою загальної теорії числень займалися Гільберт, Гедель, Пост, Каррі, Колмогоров, Марков та ін.

Числення висловлювань – перший розділ математичної логіки, що вивчає логічні операції з простими висловлюваннями, що позначаються буквами $A, B, C \ldots$, що об'єднуються в складні за допомогою логічних зв'язок. Числення висловлювань будується аксіоматично. Задається система аксіом і правила, по яких з аксіом здійснюється виведення теорем. При інтерпретації аксіоми виявляються тотожно-істинними формулами, а правила вибираються такі, які з тотожно-істинних формул забезпечують отримання тільки тотожно-істинних формул.

Слід пам'ятати, що істинні або неправдиві початкові висловлювання засобами логіки встановити не можна. Для цього потрібно звертатися до практики. Взагалі, критерієм істинності системи аксіом і будь-якої змістовної теорії є практична застосовність теорії в цілому.

3.2. Формалізації числення висловлювань

Існують різні системи числення виловлювань, які відрізняються одна від одної, зокрема тим, що вони беруть в якості вихідних даних різні набори аксіом. Так, німецький математик і логік Д. Гільберт запропонував у ролі основних зв'язок наступні: $\vee, \neg, \rightarrow$; правило виводу MP; метавизначення: $A \rightarrow B = \neg A \vee B$). Аксіоми **Теорії L_1 Гільберта, Акермана:**

$A1$:	$A \vee A \rightarrow A$	*- якщо диз'юнкція А сама із собою істинна, то і сама А істинна (**принцип тавтології**);*
$A2$:	$A \rightarrow A \vee B$	*- якщо висловлювання істинне, то і його диз'юнкція з будь-яким іншим висловлюванням також буде істинною (Рассел і Уайтхед назвали цю аксіому "**аксіомою додання**");*
$A3$:	$A \vee B \rightarrow B \vee A$	*- диз'юнкція має властивість комутативності (**закон переставляння**);*
$A4$:	$(B \rightarrow C) \rightarrow (A \vee B \rightarrow A \vee C)$	*- якщо імплікація істинна, то її члени можна диз'юнктивно зв'язати з будь-яким висловлюванням (**принцип підсумовування**; має аналог в арифметиці:* *якщо $A < B$, то $A + C < B + C$)*

Із цих аксіом за допомогою прийнятих правил Гільберт виводив нові формули.

Англійські математики і логіки Б. Расел та А. Уайтхед, зважаючи на ці п'ять аксіом, побудували своє числення висловлювань. Перші чотири аксіоми збіглися з гільбертовими, а п'ята мала такий вигляд:

$A \vee (B \vee C) \rightarrow (A \vee B) \vee C$ (*асоціативний закон*).

Але Р. Бернайс показав, що ця аксіома не є незалежною, тобто її можна вивести із перших чотирьох. Водночас він довів, що система гільбертових аксіом є незалежною.

Дана аксіоматична система не єдина в логіці. Широко поширена формальна система Кліні, що використовує 5 логічних зв'язок: $\neg, \to, \&, \vee$, правило виводу МР.

Теорія L_4 (Кліні). Схеми аксіом:

$A1: A \to (B \to A)$

$A2: \big(A \to (B \to C)\big) \to ((A \to B) \to (A \to C))$

$A3: (A \,\&\, B) \to A)$

$A4: (A \,\&\, B) \to B)$

$A5: A \to \big(B \to (A \,\&\, B)\big)$

$A6: A \to A \vee B$

$A7: B \to A \vee B$

$A8: (A \to C) \to ((B \to C) \to (A \vee B \to C))$

$A9: (A \to B) \to ((A \to \neg B) \to \neg A$

$A10: \neg\neg A \to A$

А також **Теорія L_2 (Росера)**. Основні зв'язки: $\&, \to, \neg$. Метавизначення: $A \to B = \neg(A \,\&\, \neg B)$. Схеми аксіом:

$A1: A \to (A \,\&\, A)$

$A2: (A \,\&\, B) \to A$

$A3: (A \to B) \to (\neg(B \,\&\, C) \to \neg(C \,\&\, A))$

В системі числення висловлювань, що розробив Ц. Мередіт, міститься всього дві зв'язки $\neg, \to$ і одна схема аксіом:

$\big(((A \to B) \to (\neg C \to \neg D)\big) \to E) \to ((E \to A) \to (D \to A))$

Цими самими двома зв'язками користувався Г. Фреге 1879 р. Він побудував свою систему із шести аксіом:

$A1: A \to (B \to A)$

$A2: \big(A \to (B \to C)\big) \to ((A \to B) \to (A \to C))$

$A3: \big(A \to (B \to C)\big) \to (B \to (A \to C))$

$A4: (A \to B) \to (\neg B \to A)$

$A5: \neg\neg A \to A$

$A6: A \to \neg\neg A$

Але польський математик Я. Лукасевич довів, що ця система не є незалежною, оскільки третя аксіома виводиться із кон'юнкції $A1$ і $A2$. Він залишив $A1$ і $A2$, а замість чотирьох наступних аксіом Фреге вивів одну свою: $(\neg A \to \neg B) \to (B \to A)$.

Існує і така система числення, запропонована Дж. Ніко в 1916 р., в якій всього одна зв'язка – штрих Шеффера («|») і одна схема аксіом:

$(A|(B|C)) \,\big|\, \big((D\,|(D|D))\big|\, \big((E|B)\big|\, \big((A|E)\big|(A|E)\big)\big)\big)$.

Правило виводу: $A, A|\,(B|C) \ \vdash C$.

Метавизначення: $A|B \ \equiv \ \neg A \vee \neg B$. *(Див. Мендельсон, ст. 48-52, 148-158)*

Відомі й інші системи числення висловлювань.

За правила виводу в аксіоматичних системах зазвичай беруться два: правило відділення (MP) і правило підстановки, що дозволяє замість пропозиціальної змінної підставляти будь-яку формулу в аксіоми та теореми.

Стосовно аксіоматичних побудов числення висловлювань вирішуються проблеми про їх несуперечливість, повноту та незалежність.

- Система аксіом не повинна бути суперечливою.
- Система аксіом має бути незалежною (хоча порушення цього пункту не призводить до беззмістовності теорії як, наприклад, порушення принципу несуперечливості)
- Система аксіом є семантично повною, якщо із неї можна отримати істинні на усіх наборах формули, записані засобами цього числення. Система аксіом є семантично повною (повнота у вузькому розумінні), якщо додання до неї формули, яку не можна вивести, робить її суперечливою.

Слід відзначити, що зв'язки, які використовуються в будь-якій аксіоматичній теорії, утворюють функціонально повну систему функцій у сенсі Поста.

3.3. Обчислення висловлювань. Формальна теорія L

Визначення 3.3.

1. *Символами алфавіту* теорії L є пропозиціональні літери $A, B, C, \ldots$ з індексами або без індексів, пропозиціональні зв'язки $\neg, \rightarrow$, та допоміжні символи – дужки: (та).

2. Визначення формули.

Будь-яка пропозиціональна літера є формула.

Якщо A та B є формулами, то формулами є $(\neg A), (A \rightarrow B)$.

Інших формул немає.

3. У формальній теорії L визначена нескінченна кількість аксіом, що задаються трьома схемами аксіом:

A1: $A \rightarrow (B \rightarrow A)$;

A2: $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$;

A3: $(\neg B \rightarrow \neg A) \rightarrow ((\neg B \rightarrow A) \rightarrow B)$.

Конкретні аксіоми виходять підстановкою формули замість пропозиціональної літери.

4. В теорії L визначено єдине правило виводу MP:

$A, A \rightarrow B \ \vdash B$.

Для скорочування запису формул вводяться метавизначення:

MB1: $\neg(A \rightarrow \neg B) = A \& B$.

MB2: $\neg A \rightarrow B = A \vee B$.

MB3: $(A \rightarrow B) \& (B \rightarrow A) = A \equiv B$.

3.4. Доведення та вивід у формальній теорії L

Розглянемо доведення та вивід в теорії L. Доведемо теорему $A \rightarrow A$. Оскільки єдиним правилом виводу є MP, нам треба взяти таку аксіому, щоб формула, що виводиться – $A \rightarrow A$, опинилася в конці формули. Візьмемо схему аксіоми A2, зробивши заміну B на $A \rightarrow A$ та C на A. Отримаємо

$$(A \rightarrow ((A \rightarrow A) \rightarrow A)) \rightarrow ((A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A)).$$

В схемі аксіоми А1 також замінимо В на $A \to A$, отримаємо

$A \to \big((A \to A) \to A\big)$.

Тепер до цих двох формул застосуємо правило МР, в результаті отримаємо

$\big(A \to (A \to A)\big) \to (A \to A)$.

В схемі аксіоми А1 замінимо В на А, отримаємо

$A \to (A \to A)$.

Застосовуючи до двох останніх формул правило МР, отримаємо

$A \to A$.

Доведення та вивід прийнято записувати в стовпчик, нумеруючи кожну формулу вказуючи праворуч в якості коментаря, на якій підставі формула є в цій послідовності. Нижче приводяться доведення теорем 1, 2.

Теорема 1. $\vdash A \to A$ (закон тотожності - висловлювання є логічним наслідком самого себе)

1. $(A \to ((A \to A) \to A)) \to ((A \to (A \to A)) \to (A \to A))$ А2
2. $A \to \big((A \to A) \to A\big)$ А1
3. $\big(A \to (A \to A)\big) \to (A \to A)$ МР(1, 2)
4. $A \to (A \to A)$ А1
5. $A \to A$ МР(3, 4)

Теорема 2. $\vdash (A \to A) \to A$

1. $(A \to \neg A) \to ((A \to A) \to A)$ А3
2. $\neg A \to \neg A$ Т1
3. $(\neg A \to A) \to A$ МР(1, 2)

В доведенні теореми 2 в пункті 2 ми посилалися на доведену раніше теорему 1, що фактично відповідає включенню в доведення усіх пунктів доведення теореми 1.

Побудуємо наступні виводи.

В1. Правило силогізму.

$A \to B, B \to C \vdash A \to C$

1. $A \to B$ Г1

2. $B \to C$ Г2

3. $(A \to (B \to C) \to (A \to B) \to (A \to C))$ А2

4. $(B \to C) \to (A \to (B \to C))$ А1

5. $A \to (B \to C)$ MP(2, 4)

6. $(A \to B) \to (A \to C)$ MP(3, 5)

7. $A \to C$ MP(1, 6)

В2. Правило видалення середнього посилання.

$A \to (B \to C), B \vdash A \to C$

1. $A \to (B \to C)$ Г1

2. B Г2

3. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ А2

4. $(A \to B) \to (A \to C)$ MP(1, 3)

5. $B \to (A \to B)$ А1

6. $A \to B$ MP(2, 5)

7. $A \to C$ MP(4, 6)

В3. Правило видалення крайнього посилання.

$(A \to B) \to C \vdash B \to C.$

1. $(A \to B) \to C$ Г1

2. $\mathrm{B} \to (\mathrm{A} \to \mathrm{B})$ А1

3. $B \to C$ В1(1,2)

Такі правила виводу, доведені засобами формальної теорії, називаються похідними правилами виводу.

3.5. Метатеорема про дедукцію

Метатеорема про дедукцію (МТД).

Якщо з множини формул Г і формули A виводима формула B, то з множини Г виводима формула $A \to B$, тобто якщо Г, $A \vdash B$, то Г $\vdash A \to B$.

Доведення цієї метатеореми може бути приведеним за методом математичної індукції. Нехай вивід формули B – це послідовність $B_1, B_2, \dots, B_n = B$. Доведемо наступну металему.

Металема. З $\Gamma, A \vdash B_j$ слідує, що $\Gamma \vdash A \to B_j (i = 1, \dots, n)$.

Доведення металеми.

Базис індукції. Нехай $i = 1$. Розглянемо три випадки.

а) B_1 *– аксіома.*

$\vdash B_1$

$\vdash B_1 \to (A \to B_1)$ $A1$

$\vdash A \to B_1$ за правилом МР

$\Gamma \vdash A \to B_1$ за властивістю виводимості 1.

б) $B_1 = A$, тобто B_1 є самою формулою A

$\vdash A \to A$ $T1$

Оскільки теорема виводиться з порожньої множини посилок, вона виводиться з будь-якої множини посилок, згідно властивості виводимості 1. Тому $\Gamma \vdash A \to A$, що є рівносильним до $\Gamma \vdash A \to B_1$.

в) $B_1 \in \Gamma$, тобто B_1 – гіпотеза з Г.

Доведення наводиться так само, як і в випадку а).

Крок індукції. Нехай металема виконується для всіх $k < i$. Доведемо, що вона виконується при $k = 1$. Можливі чотири випадки (доведення проводиться так само, як і для базису індукції):

а) B_i – аксіома;

б) $B_i = A$

в) $B_i \in \Gamma$

г) B_i виводиться по МР з попередніх формул, т.е. в послідовності $B_1, B_2, \dots, B_n$ є формули: B_m і $B_l = B_m \to B_i, (m < l < i)$.

Тоді по припущенню індукції справедливі виводи:

$\Gamma \vdash A \to B_m$

$\Gamma \vdash A \to B_l$, тобто $\Gamma \vdash A \to (B_m \to B_i)$.

По схемі аксіом А2

$\vdash \big(A \to (B_m \to B_i)\big) \to ((A \to B_m) \to (A \to B_i))$.

Застосовуючи двічі до останніх трьох виразів правило МР, отримаємо

$\Gamma \vdash A \to B_i$.

При $i = n$ отримаємо формулювання метатеореми про дедукцію. ◊

Справедлива метатеорема, обернена до метатеореми про дедукцію.

Обернена метатеорема про дедукцію.

Якщо існує вивід $\Gamma \vdash A \to B$, то формула B виводиться з Г і A, т.е. якщо $\Gamma \vdash A \to B$, то $\Gamma, A \vdash B$

Доведення. Нехай вивід формули $A \to B$ має вид: $B_1, \dots, B_{n-1}, A \to B$, де $B_1, \dots, B_{n-1}$ – формули з множини Г. Тоді вивід формули B р Г і A буде мати вид: $B_1, \dots, B_{n-1}, A \to B, A, B$, так як B слідує з $A \to B$ і A по правилу МР. ◊

З метатеореми про дедукцію виводяться наслідки:

Наслідок 1 (В1). **Правило силогізму**: $A \to B, B \to C \ \vdash A \to C$.

Наслідок 2 (В2). **Правило видалення середньої посилки**: $A \to (B \to C), B \vdash A \to C$.

Наслідок 3 (В3). **Правило видалення крайньої посилки**: $(A \to B) \to C \ \vdash B \to C$.

Правила силогізму і видалення середньої посилки були доведеними раніше, без використання метатеореми про дедукцію. Для порівняння наведемо доведення правила В2 з її використанням.

Користуючись оберненою метатеоремою про дедукцію, будемо будувати вивід: $(A \to B) \to C, B \vdash C$.

1.	B	Г1
2.	$(A \to B) \to C$	Г2
3.	$B \to (A \to B)$	А1
4.	$A \to B$	МР(1,3)
5.	C	МР(2,4)

Тепер, по метатеоремі про дедукцію, отримаємо вивід:

$(A \to B) \to C \vdash B \to C$.

(**В3) Правило видалення крайньої посилки**: $(A \to B) \to C \ \vdash B \to C$.

Застосовуючи обернену метатеорему про дедукцію, отримаємо:

$(A \to B) \to C, B \vdash C$.

1.	$(A \to B) \to C$	Г1
2.	B	Г2
3.	$B \to (A \to B)$	A1
4.	$A \to B$	MP(2,3)
5.	C	MP(1,4)

По метатеоремі про дедукцію, отримаємо вивід: $(A \to B) \to C \ \vdash B \to C$.

Застосування метатеореми про дедукцію (МТД) дозволяє з любого правила виводу отримати теорему. Наприклад, застосовуючи два рази МТД до правила силогізму, отримаємо теорему силогізму:

$A \to B, B \to C \vdash A \to C$	правило силогізму
$A \to B \vdash (B \to C) \to (A \to C)$	МТД
$\vdash (A \to B) \to ((B \to C) \to (A \to C))$	МТД

Обернена метатеорема про дедукцію (ОМТД) дозволяє отримувати правила виводу з теорем. Наприклад, застосовуючи два рази ОМТД до аксіоми А3, отримаємо правило виводу:

$\vdash (\neg B \to \neg A) \to ((\neg B \to A) \to B)$	А3
$\neg B \to \neg A \vdash ((\neg B \to A) \to B)$	ОМТД
$\neg B \to \neg A, \neg B \to A \vdash B$	ОМТД

Застосування метатеореми про дедукцію і наслідків з неї дозволяє спрощувати побудову виводів і доведень. Розглянемо приклади наступного застосування. (Необхідно відмітити, що наведені нижче доведення не є єдиними, можуть бути знайдені і інші доведення відповідних теорем.)

Теорема 3 (зняття подвійного заперечення). $\vdash \neg\neg A \to A$

1.	$(\neg A \to \neg\neg A) \to ((\neg A \to \neg A) \to A)$	A3
2.	$\neg A \to \neg A$	T1
3.	$(\neg A \to \neg\neg A) \to A$	B2(1,2)
4.	$\neg\neg A \to A$	B3(3)

Теорема 4 (введення подвійного заперечення). $\vdash A \to \neg\neg A$

1.	$(\neg\neg\neg A \to \neg A) \to ((\neg\neg\neg A \to A) \to \neg\neg A)$	A3
2.	$\neg\neg\neg A \to \neg A$	T3
3.	$(\neg\neg\neg A \to \neg A) \to \neg\neg A$	MP(1,2)
4.	$A \to \neg\neg A$	B3(3)

Теорема 5 (протиріччя/суперечність). $\vdash \neg A \to (A \to B)$

Побудуємо висновок: $\neg A, A \vdash B$.

1.	$\neg A$	Г1
2.	A	Г2
3.	$(\neg B \to \neg A) \to ((\neg B \to A) \to B)$	A3
4.	$\neg A \to (\neg B \to \neg A)$	A1
5.	$A \to (\neg B \to A)$	A1
6.	$\neg B \to \neg A$	MP(1,4)
7.	$\neg B \to A$	MP(2,5)
8.	$(\neg B \to A) \to B$	MP(3,6)
9.	B	MP(7,8)

Теорема 6 (контрапозиції). $\vdash (\neg A \to \neg B) \to (B \to A)$

Побудуємо висновок: $\neg A \to \neg B \vdash B \to A$

1.	$\neg A \to \neg B$	Г1
2.	$(\neg A \to \neg B) \to ((\neg A \to B) \to A)$	A3
3.	$(\neg A \to B) \to A$	MP(1,2)
4.	$B \to A$	B3(3)

Теорема 7 (контрапозиції). $\vdash (B \to A) \to (\neg A \to \neg B)$

Побудуємо висновок: $B \to A \vdash \neg A \to \neg B$

1. $B \to A$ Г1
2. $\neg\neg B \to B$ Т3
3. $A \to \neg\neg A$ Т4
4. $\neg\neg B \to A$ В1(2,1)
5. $\neg\neg B \to \neg\neg A$ В1(3,4)
6. $(\neg\neg B \to \neg\neg A) \to (\neg A \to \neg B)$ Т6
7. $\neg A \to \neg B$ МР(5,6)

Теорема 8. $\vdash A \to (\neg B \to \neg(A \to B))$

Побудуємо висновок: $A \vdash \neg B \to \neg(A \to B)$

1. $\big((A \to B) \to B\big) \to \big(\neg B \to \neg(A \to B)\big)$ Т7
2. A Г1
3. $A, A \to B \vdash B$ МР
4. $A \vdash (A \to B) \to B$ МТД(3)
5. $(A \to B) \to B$ з (2,4)
6. $\neg B \to \neg(A \to B)$ МР(1,5)

У доведенні цієї теореми фактично використовувались засоби метамови – з правила МР застосуванням МТД було отримано нове правило виводу: $A \vdash (A \to B) \to B$.

Теорема 9. $\vdash (A \to B) \to ((\neg A \to B) \to B)$

Побудуємо висновок: $A \to B, \neg A \to B \vdash B$

1. $A \to B$ Г1
2. $\neg A \to B$ Г2
3. $(A \to B) \to (\neg B \to \neg A)$ Т7
4. $(\neg A \to B) \to (\neg B \to \neg\neg A)$ Т7
5. $\neg B \to \neg A$ МР(1,3)

6. $\neg B \to \neg\neg A$ MP(2,4)

7. $(\neg B \to \neg\neg A) \to ((\neg B \to \neg A) \to B)$ A3

8. $(\neg B \to \neg A) \to B$ MP(6,7)

9. B MP(5,8)

3.6. Правила введення і видалення зв'язок

При доведенні теорем і побудові висновків можна використовувати похідні правила виведення, доведенні раніше, в тому числі правила введення і видалення зв'язок.

Таблиця 3.1.

Зв'язка	*Введення*	*Видалення*
$\to$	$\Gamma, A \vdash B \Rightarrow \Gamma \vdash A \to B$ (МТД)	$A, A \to B \vdash B$ (MP) $\Gamma_1 \vdash A, \Gamma_2 \vdash A \to B \Rightarrow \Gamma_1, \Gamma_2 \vdash B$
$\neg$	$\Gamma, A \vdash B;\ \Gamma, A \vdash \neg B \Rightarrow$ $\Rightarrow \Gamma \vdash \neg A$ (доведення від протилежного)	$A, \neg A \vdash B$ (слабке видалення заперечення) $\neg\neg A \vdash A$ (видалення подвійного заперечення)
&	$\Gamma_1 \vdash A, \Gamma_2 \vdash B \Rightarrow \Gamma_1, \Gamma_2 \vdash A\&B$	$\Gamma \vdash A\&B \Rightarrow \Gamma \vdash A$ $\Gamma \vdash A\&B \Rightarrow \Gamma \vdash B$
$\cup$	$\Gamma \vdash A \Rightarrow \Gamma \vdash A \cup B$ $\Gamma \vdash B \Rightarrow \Gamma \vdash A \cup B$	$\Gamma_1 \vdash A \cup B; \Gamma_2, A \vdash C; \Gamma_3, B \vdash C \Rightarrow$ $\Rightarrow \Gamma_1, \Gamma_2, \Gamma_3 \vdash C$
$\equiv$	$A \to B, B \to A \vdash A \equiv B$	$A \equiv B \vdash A \to B; A \equiv B \vdash B \to A$

Доведемо деякі з цих правил.

B4. $\neg$ — введення. $A \to B, A \to \neg B \vdash A$

1. $A \to B$ Г1

2. $A \to \neg B$ Г2

3. $(A \to B) \to (\neg B \to \neg A)$ T7

4. $(A \to B) \to (\neg\neg B \to \neg A)$ T7

5. $\neg B \to \neg A$ MP(1,3)

6. $\neg\neg B \to \neg A$ MP(2,4)

7. $(\neg B \to \neg A) \to ((\neg\neg B \to \neg A) \to \neg A)$ T9

8. $(\neg\neg B \to \neg A) \to \neg A$ MP(5,7)

9. $\neg A$ MP(6,8)

B5. & **– видалення 1.** $A\&B \vdash A$.

Користуючись метавизначенням MB1, отримаємо $\neg(A \to \neg B) \vdash A$, і застосуємо МТД:

$\vdash \neg(A \to \neg B) \to A$

1. $\neg A \to (A \to \neg B)$ T5

2. $(A \to \neg B) \to \neg\neg(A \to \neg B)$ T4

3. $\neg A \to \neg\neg(A \to \neg B)$ B1(1,2)

4. $\big(\neg A \to \neg\neg(A \to \neg B)\big) \to (\neg(A \to \neg B) \to A)$ T6

5. $\neg(A \to \neg B) \to A$ MP(3,4)

B6. & **– видалення 2.** $A\&B \vdash B$

$\neg(A \to \neg B) \vdash B$ MB1

1. $\neg(A \to \neg B)$ Г1

2. $\big(\neg B \to (A \to \neg B)\big) \to (\neg(A \to \neg B) \to \neg\neg B)$ T7

3. $\neg B \to (A \to \neg B)$ T5

4. $\neg(A \to \neg B) \to \neg\neg B$ MP(2,3)

5. $\neg\neg B$ MP(1,4)

6. $\neg\neg B \to B$ T3

7. B MP(5,6)

В7. **& – введення.** $A, B \vdash A\&B$

$A, B \vdash \neg(\mathrm{A} \to \neg\mathrm{B})$ МВ1

1. A Г1
2. B Г2
3. $\mathrm{A} \to (\neg\neg\mathrm{B} \to \neg(\mathrm{A} \to \neg\mathrm{B}))$ Т8
4. $\neg\neg\mathrm{B} \to \neg(\mathrm{A} \to \neg\mathrm{B})$ MP(1,3)
5. $\mathrm{B} \to \neg\neg\mathrm{B}$ Т4
6. $\neg\neg\mathrm{B}$ MP(2,5)
7. $\neg(\mathrm{A} \to \neg\mathrm{B})$ MP(4,6)

Приклади застосування правил та доведення теорем.

Приклад 3.1.Метавизначення $\vdash (\mathbf{A} \to \mathbf{B}) \to \neg\boldsymbol{A} \cup \boldsymbol{B}$

ОМТД 2 рази і МВ 2 $(\neg A \to \mathrm{B} = A \cup B)$ $A \to \mathrm{B}, \neg\neg A \vdash B$

1. $A \to \mathrm{B}$ Г1
2. $\neg\neg A$ Г2
3. $\neg\neg A \to \mathrm{A}$ Т3
4. A MP(2,3)
5. B MP(1,4)

Приклад 3.2. $\vdash \neg(\boldsymbol{A} \cup \boldsymbol{B}) \to \neg\boldsymbol{A}\&\neg\boldsymbol{B}$ **– Закон де Моргана**

ОМТД $\neg(A \cup B) \vdash \neg A\&\neg B$

МО1, МО2 $\neg(\neg A \to \mathrm{B}) \vdash (\neg A \to \neg\neg\mathrm{B})$

Доведемо допоміжну лему

ЛЕМА $\vdash (\neg A \to \neg\neg\mathrm{B}) \to (\neg A \to \mathrm{B})$

ОМТД $(\neg A \to \neg\neg\mathrm{B}), \mathrm{A} \vdash B$

1. $\neg A \to \neg\neg\mathrm{B}$ Г1
2. $\neg A$ Г2
3. $\neg\neg\mathrm{B}$ MP(1,2)
4. $\neg\neg\mathrm{B} \to \mathrm{B}$ Т3
5. B MP(3,4)

Продовжимо доведення початкової теореми

1. $\neg(\neg A \to \text{B})$ Г1
2. $\vdash (\neg A \to \neg\neg\text{B}) \to (\neg A \to \text{B})$ Лема
3. $\big((\neg A \to \neg\neg\text{B}) \to (\neg A \to \text{B})\big) \to \big(\neg(\neg A \to \text{B}) \to \neg(\neg A \to \neg\neg\text{B})\big)$ Т7
4. $\big(\neg(\neg A \to \text{B}) \to \neg(\neg A \to \neg\neg\text{B})\big)$ МР(2,3)
5. $\neg(\neg A \to \neg\neg\text{B})$ МР(1,4)

Приклад 3.3. $\vdash \boldsymbol{A} \cup (\boldsymbol{B}\&\mathbf{C}) \to (\boldsymbol{A} \cup \boldsymbol{B})\&(\boldsymbol{A} \cup \boldsymbol{C})$ – **Дистрибутивний закон**

ЛЕМА $\vdash (\boldsymbol{a} \to \mathbf{b})\&(\boldsymbol{a} \to \mathbf{c}) \to (\boldsymbol{a} \to \mathbf{b}\&\mathbf{c})$**,** де консеквент - схема початкової теореми

Для повного доведення необхідно

- Довести лему
- Довести окремо положення $(a \to \text{b})$ і $(a \to \text{c})$ в термінах початкової теореми, тобто

$\vdash A \cup (B\&\text{C}) \to (A \cup B)$ і $\vdash A \cup (B\&\text{C}) \to (A \cup C)$

- З леми і двох доведених теорем отримати докази початкової шляхом застосування МР.

Приклад 3.4. $\vdash ((\boldsymbol{A} \to \mathbf{B}) \to \mathbf{A}) \to \mathbf{A}$ **– Закон Пірса**

ОМТД $(A \to \text{B}) \to \text{A} \vdash A$;

1. $(A \to \text{B}) \to \text{A}$ Г1
2. $\vdash \neg\text{A} \to (A \to \text{B})$ Т5
3. $\neg\text{A} \to \text{A}$ В1 (1,2)
4. $(\neg\text{A} \to \text{A}) \to \text{A}$ Т2
5. A МР 3,4

Застосування засобів формальної теорії L при перевірці висновків у логічних задачах природної мови

Приклад 3.5. Після формалізації речення природньої мови, наприклад, коти або собаки на пляж не допускаються, отримаємо такий формальний запис:

$K \vee C \to B$, *де К – коти, С – собаки, В – не допускаються на пляж.*

Перетворимо за еквівалентними формулами вираз: $K \vee C \to B = \neg(K \vee C) \vee B =$

$= (\neg K \& \neg C) \vee B = (\neg K \vee B) \& (\neg C \vee B) = (K \to B) \& (C \to B)$, що надає нам можливість розбити гіпотезу на дві: $K \to B$ *та* $C \to B$.

Такий же результат ми отримаємо, застосовуючи правило видалення крайньої посилки (В3): $K \vee C \to B = (\neg K \to C) \to B$. Після В3 : $C \to B$.

Якщо, $K \vee C \to B = (\neg K \to C) \to B$ *(1)* та візьмемо правило контрапозиції у такому вигляді $(\neg C \to K) \to (\neg K \to C)$ *(2)*, застосуємо силогізм 1-2: $(\neg C \to K) \to B$.

Далі використання В3 дасть нам $K \to B$.

Тобто в формальному виводі можна записати:

1. $K \vee C \to B$ *г1*
2. $K \to B$
3. $C \to B$ пп.2 та 3 отриманні за допомогою еквівалентних перетворювань та видалення &

Або:

1. $K \vee C \to B$ *г1*
2. $K \to B$
3. $C \to B$ пп.2 та 3 отриманні за допомогою видалення крайньої посилки

Приклад 3.6. В задачах, де для отримання висновку на поверхні лежить використання правила доведення до абсурду, можна застосувати такий елегантний хід, як В1 (силогізм)+Т2 теорії L.

1) Нехай в процесі побудови формального висновку отримано дві гіпотези $A\to B$ та $A\to \neg B$. Зрозуміло, що за правилом доведення до абсурду ми повинні отримати $\neg A$. Розглянемо інший шлях:

1. $A\to B$ г1
2. $A\to \neg B$ г2
3. $(A\to \neg B) \to (B\to \neg A)$ *правило контрапозиції*
4. $B\to \neg A$ *MP 2,3*
5. $A\to \neg A$ *B1-1,4*
6. $\neg\neg A\to A$ *T4*
7. $\neg\neg A\to \neg A$ *B1-5,6*
8. $(\neg\neg A\to \neg A) \to \neg A$ *T2*
9. $\neg A$ *MP 7,8*

2) За умовою вихідних гіпотез $B\to A$ та $\neg B\to A$ проведемо вивід:

1. $B\to A$ г1
2. $\neg B \to A$ г2
3. $(\neg B \to A) \to (\neg A \to B)$ *правило контрапозиції*
4. $\neg A \to B$ *MP 2,3*
5. $\neg A\to A$ *B1-1,4*
6. $(\neg A\to A) \to A$ *T2*
7. A *MP 5,6*

При доведенні теорем в теорії L в аналогічних ситуаціях при наявності гіпотез таких, як в п.1 та п.2, підходить для застосування правило доведення до абсурду, що формалізується у вигляді А3 або Т9. Але вищенаведений підхід через силогізм та Т2 не змушує згадувати різницю в запису методу абсурду через А3 або Т9.

3.7. Властивості формальної теорії L

Для дослідження властивостей формальної теорії, необхідно побудувати її модель (інтерпретацію), в якій кожному реченню в формальній теорії повинно відповідати речення моделі.

Поставимо у відповідність кожній літері обчислення L пропозиціональну літеру, кожній формулі — формулу, кожній теоремі — тавтологію алгебри висловлювань. Тоді, ми отримаємо взаємно однозначну відповідність, яка доводить, що алгебра висловлювань являється моделлю теорії L.

Визначення 3.4. Формальна теорія є *повною відносно моделі*, якщо кожній теоремі відповідає тотожно істинна формула моделі. *Модель* (або *інтерпретація*) *адекватна* формальній теорії, якщо кожна формула моделі, яка є істинною, є теоремою теорії.

Теорема 3.1. Кожна теорема теорії L є тавтологією алгебри висловлювань.

Доведення. Схемам аксіом A1, A2, A3 відповідають тавтології алгебри висловлювань. Для перевірки цього положення достатньо побудувати їх таблиці істинності. Правило MP зберігає властивість тавтологічності опираючись на теорему 2.1 про тавтології. Так, як будь-яка теорема виводиться із аксіом за допомогою правила MP, їй також буде відповідати тавтологія алгебри висловлювань.◊

З цього слідує, що формальна теорія L є *повною відносно алгебри висловлювань*.

Теорема 3.2 (**Метатеорема адекватності**). Кожна тавтологія алгебри висловлювань є теоремою формальної теорії L.

Доведення базується на наступній металемі.

Металема. Нехай формула A алгебри висловлювань залежна від пропозиціональних букв $B_1, B_2, \dots, B_k$. Тоді, кожному рядку таблиці істинності цієї формули відповідає висновок формальної теорії L виду $B_1', \dots, B_k', \vdash A'$, де

$B' = B$, якщо $|B| = T$, і $B' = \neg B$, якщо $|B| = F$, $\neg A = A$ або $\neg A$, якщо $|A| =$ T або F відповідно.

Сенс даної леми полягає в наступному. Якщо дана формула, наприклад, $E = B \to ((A \to B) \to A)$, то, побудувавши її таблицю істинності (табл. 3.2), можна визначити, з яких посилок вона виводиться.

Таблиця.11.2.

A	B	$(A \to B) \to A$	$B \to ((A \to B) \to A)$
F	F	F	T
F	T	F	F
T	F	T	T
T	T	T	T

За металемою, для даної формули можна побудувати чотири висновки: $\neg A, \neg B \vdash E$; $\neg A, B \vdash \neg E$; $A, \neg B \vdash E$; $A, B \vdash E$.

Доведення металеми.

Доведення проводиться індукцією по числу зв'язок n в формулі A.

1. *Базис індукції*. Нехай $n = 0$, тоді $A = B$, тобто A представляє собою просто пропозиціональну букву B. Тоді формула A може приймати одне з двох значень, T або F, кожному з яких відповідає висновок: $B \vdash B$ и $\neg B \vdash \neg B$.

2. Крок *індукції*. Припустимо, число зв'язок в формулі A дорівнює m і при цьому виконується металема. Доведемо, що металема виконується, якщо $n =$ $= m + 1$.

1 випадок. Формула A отримана за допомогою зв'язки заперечення: $A =$ $= \neg B$, де B містить m зв'язок. Можливі наступні випадки:

а) $|B| = T$, тоді $|\neg B| = F$, тобто $|A| = F$. Необхідно довести, що існує висновок:

$B'_1, \ldots, B'_n \vdash \neg A$.

Згідно індуктивному припущенню, існує висновок: $B'_1, \ldots, B'_n \vdash B$. Згідно теореми 4, $B \vdash \neg\neg B$, тоді по правилу силогізму $B'_1, \ldots, B'_n \vdash \neg\neg B$. Но $\neg\neg B =$ $\neg A$, оскільки $A = \neg B$, отже, $B'_1, \ldots, B'_n \vdash \neg A$.

b) $|B| = F,\ |\neg B| = T,\ |A| = T$. Необхідно довести, що існує висновок $B_1', \dots, B_n' \vdash A$, або, що те ж саме, $B_1', \dots, B_n' \vdash \neg B$.

Так, як $|B| = F$, то $\neg B = B'$. Висновок $B_1', \dots, B_n' \vdash B'$ існує згідно індуктивного припущення, а значить, існує и $B_1', \dots, B_n' \vdash \neg B$.

2 випадок. Формула A сформована за допомогою логічної зв'язки імплікації: $A = B \to C$**,** де A и C містять не більше m зв'язок. За індуктивним припущенням існують висновки:

$B_1', \dots, B_n' \ \vdash B'$,

$B_1', \dots, B_n' \vdash C'$.

Розглянемо істинносні розподіли.

a) $|C| = T, |A| = T$.

Треба довести, що існує висновок $B_1', \dots, B_n' \vdash A$, т.е. $B_1', \dots, B_n' \vdash B \to C$.

За індуктивним припущенням існує висновок $B_1', \dots, B_n' \vdash C$. Тоді, з цього висновку і аксіоми A_1: $C \to (B \to C)$ по MP отримуємо: $B_1', \dots, B_n' \vdash B \to C$.

b) $|B| = F, |A| = T$.

Висновок $B_1', \dots, B_n' \vdash \neg B$ існує за індуктивним припущенням. Тоді з цього висновку і теореми $T5$: $\vdash\ \neg B \to\ (B \to C)$ за правилом MP отримуємо: $B_1', \dots, B_n' \vdash B \to C$.

c) $|C| = F, |B| = T, |A| = F$. Довести, що існує висновок $B_1', \dots, B_n' \vdash \neg A$, тобто

$B_1', \dots, B_n' \vdash \neg(B \to C)$. Побудуємо цей висновок.

1. $B_1', \dots, B_n' \vdash \neg C$ (за індуктивним припущенням)
2. $B_1', \dots, B_n' \vdash B$ (за індуктивним припущенням)
3. $\vdash B \to (\neg C \to \neg(B \to C))$ $T8$
4. $B_1', \dots, B_n' \vdash \neg C \to \neg(B \to C) MP$(2,3)
5. B_1', $A = \neg\neg B, \dots, B_n' \vdash \neg(B \to C)$ MP(1,4)

Металема доведена.◊

Доведення метатеореми адекватності.

Нехай А - тавтологія, що залежить від пропозиціональних букв $B_1, \dots, B_n$. Згідно металеммі, для кожного істинносного розподілу пропозиціональних букв існують висновки $B_1', \dots, B_n \vdash A$ (А' збігається з А істинно в кожному рядку таблиці істинності). У таблиці істинності є два рядки, які розрізняються лише значенням істинності B_n. Для цих рядків існують висновки:

$B_1', \dots, B_n \vdash A$ *, де* $|B_n| = T$ та $B_1', \dots, \neg B_n \vdash A$*, де* $|B_n| = F$

Застосуємо до них метатеорему про дедукцію. отримаємо:

$B_1', \dots, B_{n-1} \vdash B_n \to A,$

$B_1', \dots, B_{n-1} \vdash \neg B_n \to A.$

Візьмемо теорему Т9: $\vdash (B_n \to A) \to \big((\neg B_n \to A) \to A\big)$.

Застосовуючи двічі правило МР, отримаємо, висновок. $B_1', \dots, B_{n-1} \vdash A$

Аналогічним чином ми можемо виключити всі змінні, і за n кроків отримаємо $\vdash A$. ◊

Теорема адекватності встановлює взаємно однозначну відповідність між теоремами теорії L і тавтологіями алгебри висловлювань. Звідси випливає, що алгебра висловлювань адекватна формальнії теорії L.

Розглянемо властивості формальної теорії L.

Визначення 3.5. Формальна теорія несуперечлива, якщо не існують такої формули А, щоб A и ¬А одночасно були теоремами теорії.

Теорема 3.3. Теорія L несуперечлива.

Доведення. Кожна теорема теорії L є тавтологією алгебри висловлювань. Заперечення формули, що являється тавтологією, тавтологією не є. Отже, заперечення теореми теоремою не є.◊

Повнота теорії розуміється у вузькому і широкому сенсі.

Визначення 3.6. Теорія є повною (у вузькому сенсі), якщо додавання будь-якої недовідної в цій теорії формули робить її суперечливою. Повнота в широкому сенсі означає, що кожну формулу можна довести або спростувати, тобто або $\vdash A$, або $\vdash \neg A$.

Теорема 3.4. Теорія L неповна в широкому сенсі.

*Доведення***.** Дійсно, не кожна формула або її заперечення є теоремами теорії L. Якщо взяти нейтральну формулу алгебри висловлювань, то її заперечення також є нейтральною формулою, тобто ні сама формула, ні її заперечення не є тавтологією алгебри висловлювань. Тому ці формули не є теоремами обчислення L. ◊

Теорема 3.5. Теорія L повна у вузькому сенсі.

Для доведення теореми потрібно показати, що теорія L стає суперечливою при додаванні до її системі аксіом будь-якої недовідної в цій теорії формули.

Доведення. Теорія L має три схеми аксіом A1, A2, A3 і правило виведення MP. Побудуємо нову теорію L', додавши до системи аксіом L формулу A, яка не є тавтологією алгебри висловлювань. Тоді формула A приймає хоча б одне хибне значення на деякій інтерпретації. Значить, якщо A представлена кон'юнктивною нормальною формою, то ця форма повинна містити хоча б одну елементарну диз'юнкцію δ, що не містить ні одну змінну разом з її запереченням: $\delta = B'_1 \vee B'_2 \vee \ldots \vee B'_n$, де $B'_i = \neg B_i$, якщо $\left|B_i\right| = T$, і $B'_i = B_i$, якщо $\left|B_i\right| = F$. В елементарній диз'юнкції δ замінимо кожне входження пропозиціональної букви B'_i на B, якщо $|B_i| = F$, і на $\neg B$, якщо $|B_i| = $ T. Отримаємо: $\delta' = B \vee \neg\neg B \vee \ldots \vee B \vee \neg\neg B = $ B. Зробимо іншу заміну: замінимо B'_i на $\neg B$, якщо $|B_i| = F$, і на B, якщо $|B_i| = $ T. Отримаємо: $\delta'' = \neg B \vee \neg B \vee \ldots \vee \neg B = \neg$B. У новій теорії L' формула A - аксіома, тобто $\vdash A$. Оскільки A представлена у вигляді ДКНФ, то за правилом видалення & будь-яка диз'юнкція δ кон'юнктивної нормальної форми також доказова, а оскільки диз'юнкція δ представлена у вигляді δ′, так і у вигляді δ″, то у L' має місце $\vdash \delta'$ и $\vdash \delta''$, а це означає, що $\vdash \neg B$ та $\vdash B$, тобто теорія L' суперечлива.◊

Визначення 3.7. Формальна теорія називається *вирішуваною*, якщо існує ефективна процедура, що дозволяє за кінцеве число кроків визначити, є довільна формула теоремою чи ні.

Теорія L вирішувана, так як кожної теоремі теорії відповідає тавтологія, а для будь-якої тавтології можна побудувати таблицю істинності.

Визначення 3.8. *Система аксіом є незалежною,* якщо жодна з аксіом не може бути виведена з інших.

Теорема 3.6. Схеми аксіом A1, A2, A3 в теорії L незалежні.

*Доведення***.** Доведемо незалежність A1. Для цього необхідно побудувати таку несуперечливу модель, в котрій виконуються всі аксіоми, крім першої. Побудуємо модель в тризначній логіці, де істинності значення операцій ¬ і → визначені в табл. 3.3, 3.4.

Формулу A будемо вважати виділеною, якщо вона завжди приймає значення 0. Неважко показати, що правило MP зберігає властивість виділенності. Можна показати (побудовою таблиць істинності), що схеми аксіом A2, A3 є виділеними в даній моделі. Отже, виділеною є і всяка формула, виведена з A2, A3 за допомогою правила MP. Однак формула A1 невиділенна, – для доказу цього достатньо знайти один набір, на якому значення A1 відмінне від 0, наприклад: $1 \to (2 \to 1) = 1 \to 0 = 2$.

Аналогічно можна довести незалежність A2, побудувавши іншу тризначну модель, де виділеними будуть аксіоми A1, A3 (детальне доведення можна знайти в [Мендельсон, 1976]). Щоб показати незалежність A3 досить перевизначити заперечення так, щоб ¬x = x (тотожна операція). Тоді A1 і A2 як і раніше будуть тавтологіями, а A3 вже не буде тавтологією.

Можна спробувати побудувати свою аксіоматичну теорію, вибравши в якості аксіом закони класичної логіки: закон тотожності, закон суперечності, закон виключення третього, дистрибутивні закони, закон монотонності достовірних міркувань, закон силогізму, закон контрапозиції тощо.

Таблица.3.3.

A	$\neg A$
0	1
1	1
2	0

Таблица.3.4.

A	B	$A \to B$
0	0	0
1	0	2
2	0	0
0	1	2
1	1	2
2	1	0
0	2	2
1	2	0
2	2	0

З вихідних аксіом за певними правилами можна виводити теореми – інші тотожно-істинні формули.

Хоча обчислення висловлювань має велике значення, це тільки перша сходинка математичної логіки. Обчислення висловлювань, що оперує з висловлюваннями, які навіть не розчленовуються на суб'єкт і предикат, тобто на підмет і присудок, має вкрай обмежені ресурси для дослідження найелементарніших суджень, що зустрічаються в науковій і практичній діяльності. Обчислення висловлювань достатньо лише для вираження тих логічних зв’язків, в яких висловлювання виступають як нерозривне ціле. Марною була спроба відшукати формальне подання для логічного зв'язку, яка виражається наступним силогізмом: *Всі люди смертні. Сократ - людина. Отже, Сократ - смертний.* Пояснюється це тим, що в подібних висновках мова йде не тільки про висловлювання як про ціле, а про висловлювання, в яких істотну роль відіграє внутрішня логічна структура висловлювань, що виражається в тому, що в кожному з трьох висловлювань, наявних в силогізмі, є суб'єкт і предикат. Наступною сходинкою в дослідженні логічних зв’язків між висловлюваннями є обчислення предикатів, область математичної логіки, в якій висловлювання розчленовуються.

Аксіоматичний метод має найважливіше значення для математики і математичної логіки. Аксіоматичний метод полегшує організацію і систематизацію наукового знання, дозволяє швидше виявити внутрішній логічний зв'язок між окремими розділами теорії, чітко виокремлює вихідні положення і положення, що отримуються з аксіом, привчає до точності і строгості суджень, являє найцінніший інструмент наукового дослідження, пошуку нових математичних закономірностей.

Але, застосовуючи аксіоматичний метод побудови наукової теорії, треба мати на увазі, що він не може абсолютизуватися. Так, встановлення факту несуперечності має велике значення в процесі з'ясування істинності цієї системи, тому що наявність протиріччя руйнує систему. Але встановлення несуперечності - це тільки одна з вимог, що пред'являються до теорії. Робота К. Геделя в цьому напрямку виявила неможливість повної формалізації мислення, а, отже, і відому обмеженість аксіоматичного методу. У теоремі про неповноту формальних систем Гедель довів, що будь-яка досить багата несуперечлива формальна система неодмінно неповна, тому що в ній можна побудувати деяку формулу, яку буде важко вирішити в цій системі.

Якщо в концепції гільбертонівського аксіоматичного методу закон виключення третього виступає в якості логічної аксіоми, то в інтуїционістській логіці - заперечується застосовність його в операціях з нескінченними множинами, а метод моделей (інтерпретації), широко використовуваний в процесі з'ясування несуперечності, для конструктивної логіки не підходить, так як побудована формальна система вважається коректною тільки, якщо зазначений спосіб потенційно здійсненної побудови об'єктів формальної системи.

Істинність формальної системи не можна визначити, якщо не знайдена відповідна їй інтерпретація, тобто якщо ця формальна система (ФС) і її вихідні положення не поширені на якусь змістовну систему. Наприклад, обчислення висловлювань є істинною ФС, тому що знайшла підтвердження в інтерпретації релейно-контактних схем. Зазвичай ФС створюється як ідеалізований образ чогось іншого, і це інше можна назвати теорією по відношенні до чого ця ФС набуває сенсу. ФС є формалізацію деякої теорії, утворюючи синтаксис останньої, а теорія - інтерпретацію даної ФС. Дійсно, правильно побудовані слова ФС отримують сенс в рамках формалізованої теорії, а будь-якому висловлюванню, який має сенс у цій теорії, відповідає певний клас правильно побудованих слів в ФС.

Питання для самоконтролю до розділу 3

1. Визначення формальної теорії. Виведення, доведення.
2. Властивості виведення.
3. Аксіоми та правила формальної теорії *L*.
4. Теореми теорії *L*.
5. Похідні правила виводу.
6. МТД
7. Метатеорема адекватності.
8. Властивості системи аксіом.
9. Які формальні теорії числення висловлювань ще відомі.
10. Практичне застосування формальних теорій.

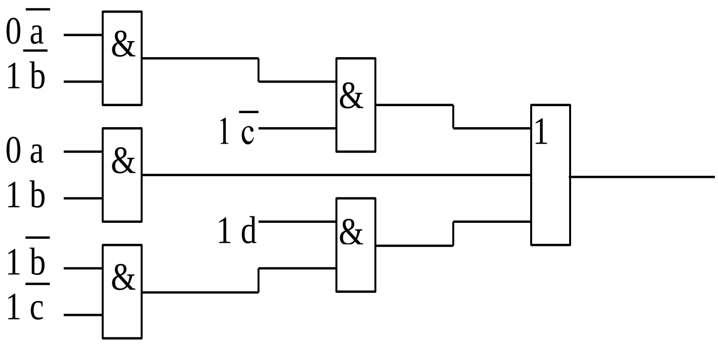
4. ТЕОРІЯ АБСТРАКТНИХ АВТОМАТІВ

Теорія автоматів відноситься до числа ключових розділів сучасної кібернетики. Вона безпосередньо пов'язана з математичною логікою, теорією алгоритмів і теорією формальних граматик. Універсальність і відносна простота обумовили розповсюджене використання автоматних моделей на практиці. Теорія автоматів, наприклад, успішно використовується при побудові вузлів цифрових обчислювальних машин, при побудові програм, і, зокрема, лексичних аналізаторів в трансляторах.

4.1. Тривіальні автомати

Відомо, що в якості інтерпретацій булевих функцій застосовують електричні та комбінаційні схеми. У комбінаційній схемі вихідний сигнал у будь-який момент часу залежить від вхідного сигналу у той же момент часу. Комбінаційна схема приймає лише один стан. Комбінаційні схеми представляють собою автомат без пам'яті, тобто тривіальний автомат.

Комбінаційні схеми повністю описуються булевими функціями (однією або системою). Наприклад:



f(a,b,c,d)=¬a¬b¬c ∨ ab ∨ bcd=0   при a=0, b=1, c=1, d=1 .

При синтезі комбінаційної схеми вихідний стан задається або у вигляді *системи булевих функцій*, або як *таблиця істинності*, які визначають поведінку системи. Мета синтезу полягає в побудові структури схеми на основі даної системи логічних елементів так, щоб її поведінка задовольняла даній системі булевих функцій.

У подібних контактних і логічних схемах значення змінних виходу визначаються тільки комбінацією значень змінних на входах в даний момент часу. Тому їх і називають комбінаційними схемами. У більш широкому випадку змінні виходу можуть залежати від значень вхідних змінних не тільки у даний момент часу, але й від попередніх значень. Інакше кажучи, значення змінних виходу визначаються послідовністю значень вхідних змінних, у зв'язку з чим схеми з такими властивостями називають послідовносними. Якщо змінні входу та виходу приймають значення кінцевих алфавітів, то обидва типи схем об'єднуються в назву *скінченних автоматів*.

В реальних умовах сигнали представляють собою неперервні функції часу, тому для надійного розрізняння сигналів потребується, щоб нові значення на входах з'являлись після завершення перехідних процесів, зв'язаних з попередніми значеннями. В процесі розгляду логічної структури автоматів як правило відволікаються від існування цих процесів і вважають, що аргументи змінюються не неперервно, а миттєво у деякі моменти часу (*такти*). Інтервали між тактами можуть бути відмінними, але без втрати спільності їх можна вважати рівними.

Окрім змінних входу та виходу можна виділити деяку множину проміжних аргументів, які пов'язані з внутрішньою структурою автомата. У комбінаційних схемах проміжні аргументи не беруть участь в співвідношенні вхід-вихід. Навпаки, функції виходу послідовних схем у якості своїх аргументв, окрім змінних входу обов'язково містять деяку множину проміжних аргументів, що характеризують стан схеми. Набір усіх можливих станів, які мають місце в цій схемі називаться *множиною станів*.

Таким чином, стан скінченного автомату у будь-який тактовий момент характеризуться значеннями такої множини змінних, яка разом із заданими значеннями змінних входу дозволяє визначити змінні виходу в даний тактовий момент і стан в наступний тактовий момент.

Очевидно, що послідовні схеми повинні володіти здатністю зберігати попередній стан до наступного такту, у зв'язку з чим їх називають також автоматами з пам'яттю. *Тривіальні автомати* мають тільки один стан.

4.2. Абстрактні скінченні автомати

У техніці з поняттям автомату пов'язаний деякий пристрій, здатний виконувати деякі функції без втручання людини або з обмеженням його участі. Проте таке розуміння автомату є занадто вузьким. У загальному розумінні скінченний автомат – це математична модель, що відображає фізичні або абстрактні явища найрізноманітнішої природи.

Отже, абстрактними автоматами ми будемо називати дискретні перетворювачі інформації, що видають деякий сигнал виходу (літеру алфавіту виходу) у відповідь на деякий вхідний сигнал (літеру вхідного алфавіту). Внутрішня структура автомату не розглядається.

Таким чином, абстрактний автомат задається трьома множинами:

- вхідний алфавіт *X* (множина вхідних символів)
- вихідний алфавіт *У* (множина вихідних символів)

• множина внутрішніх станів A (множина станів).

Автомат працює в дискретному часі, послідовність якого ототожнюють з натуральними числами: $t=0,1,2\ldots\ldots$

У кожен момент часу t автомат знаходиться в одному із станів $a = a(t)$ із множини A. Стан в момент $t = 0{:}a_1\ a(0)$ називається початковим станом. Початковий стан автомату залишається незмінним за будь-яких експериментів над ним. Автомати із заданим початковим станом називаються *ініціальними*.

У кожен момент часу t, починаючи з $t = 1$, на вхід автомату в якості вхідного сигналу надходить одна з літер алфавіту $X = X(t)$.

Кінцеві упорядковані послідовності вхідних сигналів $X(1), X(2)\ ..\ X(k)$ називаються *вхідними словами* автомату.

На вхід автомату може подаватися будь-яке вхідне слово із деякої множини допустимих вхідних слів.

Будь-яке допустиме вхідне слово, подане на вхід, викликає появу вихідного слова $Y(1),Y(2)\ldots Y(k)$.

Вихідне слово – також деяка впорядкована послідовність вихідних сигналів, що має ту саму довжину, що і вхідне слово. Таким чином, автомат відображає вхідні слова в вихідні.

Це відображення позначимо ϕ, таке, що однозначно визначається заданням двох функцій σ та λ , що називаються функціями переходів і функцій виходів.

Функція переходів σ визначає стан $a(t)$ автомата у будь-який момент дискретного часу t по вхідному сигналу $x(t)$ і стану у попередній момент часу $a(t-1){:}\ \sigma{:}A\times X\rightarrow A \quad a(t+1)=\sigma(a(t),x(t))$.

Функція виходів визначає залежність вихідного сигналу від змінних: $\lambda{:}A\times X\rightarrow Y \quad y(t)=\lambda(a(t),x(t))$.

Задаючи вхідне слово $p = x(1), x(2),\ldots, x(k)$ і початковий стан автомату $a(0)$ за допомогою σ та λ можна визначити вихідне слово $q=y(1)y(2)\ldots y(k)$: $q=\phi(p)$.

Автомат називається *повністю визначеним*, якщо він визначений на усіх парах (A,X) и частковим в іншому випадку.

Таким чином скінченний автомат однозначно визначається початковим станом $a(0)$ і функціями переходів і виходів, і є п’ятіркою $A=\{X, Y, A, \lambda, \sigma\}$.

Абстрактна теорія автоматів виділяє два основні способи використання автоматів: перетворення вхідної послідовності символів у вихідну (синтез дискретних пристроїв) і перевірка правильності вхідної послідовності символів (синтез програмних аналізаторів).

4.3. Способи завдання автоматів. Автомати Мілі і Мура

Закони функціонування автоматів задаються рівняннями:

а) для автоматів першого роду (автоматів Мілі)

$$a(t) = \sigma(a(t-1), x(t)),$$
$$y(t) = \lambda(a(t-1), x(t));$$

б) для автоматів другого роду

$$a(t) = \sigma(a(t-1), x(t)),$$
$$y(t) = \lambda(a(t), x(t));$$

в) для правильних автоматів другого роду (автоматів Мура) вихідні сигнали залежать тільки від стану автомату

$$a(t) = \sigma(a(t-1), x(t)),$$
$$y(t) = \lambda(a(t)).$$

Опис роботи автоматів може бути задано за допомогою таблиць переходів и виходів. Наприклад, автомат Мілі:

$A=\sigma(a,x)$ $y=\lambda(a,x)$

	A_1	A_2	A_3	A_4			A_1	A_2	A_3	A_4
X_1	A_4	A_3	A_2	A_1		X_1	Y_1	Y_2	Y_3	Y_4
X_2	A_1	A_4	A_3	A_2		X_2	Y_4	Y_1	Y_2	Y_3
X_3	A_2	A_1	A_4	A_3		X_3	Y_3	Y_4	Y_1	Y_2

Таким чином, якщо у кінцевий момент часу поданий вхідний сигнал $X_{2,}$ а автомат знаходиться у стані $A_{4,}$ то він переходить у стан A_2. У стані A_2 при вхідному сигналі X_2 на виході отримуємо Y_1. І так далі.

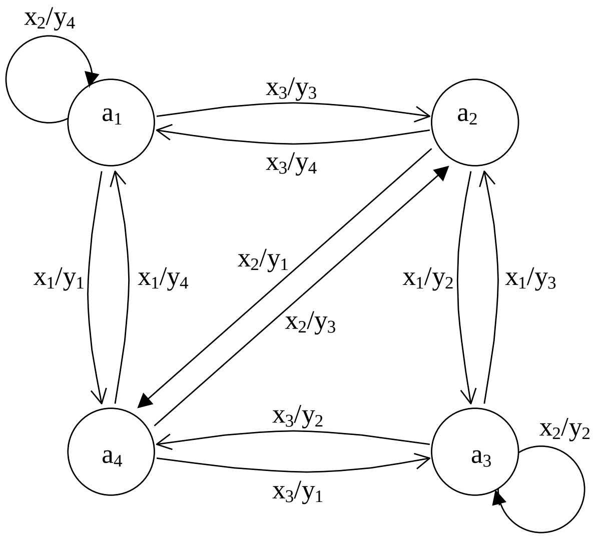
Для задання автомату Мура зручно користуватися відміченою таблицею переходів, наприклад:

	y_3	y_2	y_1
	a_1	a_2	a_3
x_1	a_3	a_2	a_1
x_2	a_1	a_3	a_2
x_3	a_2	a_1	a_3

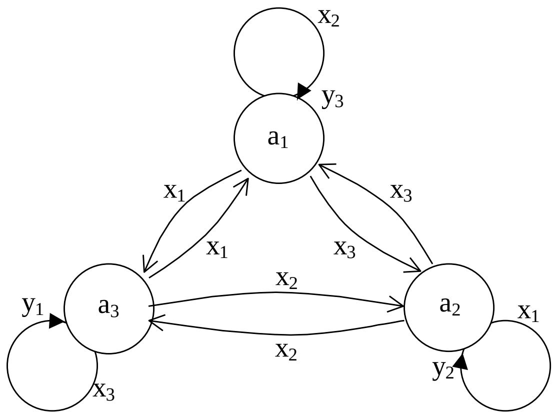
Якщо подати на вхід x1, а автомат знаходився в стані a3, то він перейде в стан a1, а на виході буде y1. Стовбці таблиці відмічаються вихідним сигналом, що є придатний даному стану.

Автомати також представляють у вигляді графів, де вершини відповідають станам автомату, а зазначені дуги – переходам із стану в стан при відповідному вхідному сигналі (відмічається - вхідний/вихідний сигнал).

Автомат Мілі:



Автомат Мура:



Автомат називається *частковим*, якщо деякі комбінації "стан – вхідний сигнал" не можуть виникнути в реальних умовах. При цьому у графі автомату з'являються стани, з яких визначені виходи не для усіх вхідних сигналів (тобто присутні не усе стрілки), а в таблицях переходів і виходів (і в зазначеній таблиці переходів) мають місце порожні клітинки. Для виконання еквівалентних перетворень, як і для структурного синтезу, необхідно довизначити частковий автомат. Перехід і виходи зазвичай довизначають, виходячи з міркувань зручності мінімізації.

4.4. Особливості теорії автоматів

Поряд з поняттям кінцевого автомата розглядаються різні його узагальнення і модифікації, що відображають ті чи інші особливості реальних пристроїв. Для кінцевого автомата існуючі модифікації можна розбити на наступні три основні групи. До першої групи належать автомати, у яких деякі з алфавітів вхідний, станів або вихідний нескінченні, в зв'язку з чим такі автомати називаються нескінченними. До другої групи належать автомати, у яких замість вихідний і перехідною функцій допускаються довільні відношення або випадкові функції. Такі – часткові, недетерміновані, ймовірнісні та інші автомати. До третьої групи відносяться автомати зі специфічними множинами вхідних об'єктів. Такі, наприклад, автомати зі змінною структурою. Існують автомати, що належать одночасно різним групам. Поряд з цим велику роль відіграють спеціальні підкласи кінцевих автоматів, наприклад, автомати без пам'яті. Крім того, використання понять і методів з інших розділів математики також призводить до появи специфічних класів автоматів і пов'язаних з ними задач. Наприклад, при застосуванні алгебраїчних засобів виникають поняття автоматів над термами; питання теорії кодування породжують поняття самоналагоджувальних, оборотних автоматів та інші.

Теорія автоматів - це розділ теорії керуючих систем, що вивчає математичні моделі перетворювачів дискретної інформації, звані автоматами. З певної точки зору такими перетворювачами є як реальні пристрої (обчислювальні машини, автомати, живі організми і т.д.), так і абстрактні системи (наприклад, формальна система, аксіоматичні теорії і т.д.). Характерною особливістю такого опису є дискретність відповідних математичних моделей і кінцевість областей значень їх параметрів, що приводить до поняття кінцевого автомата. Найбільш тісно теорія автоматів пов'язана з теорією алгоритмів.

Більшість задач теорії автоматів - загальні для основних видів керуючих систем. До них відносяться задачі аналізу і синтезу автоматів, задачі повноти, мінімізації, еквівалентних перетворень автоматів та інші.

Задача аналізу полягає в тому, щоб по заданому автомату описати його поведінку або за неповними даними про автомат і його функціонування встановити ті чи інші його властивості.

Задача синтезу автоматів полягає в побудові автомата з наперед заданою поведінкою або функціонуванням.

Задача повноти полягає у з'ясуванні, чи володіє множина автоматів властивістю повноти, тобто збігається множина всіх автоматів, які виходять шляхом кінцевого числа застосувань деяких операцій до автоматів, із заданої підмножини автоматів.

Задача еквівалентних перетворень в загальному вигляді полягає в тому, щоб знайти систему правил перетворень (так звану повну систему правил) автоматів, які задовольняють певним умовам і дозволяють перетворити довільний автомат в будь-який еквівалентний йому автомат. Два автомата еквівалентні, якщо вони мають однакову поведінку автомата; поведінка автомата - математичне поняття, яке описує взаємодію автомата з зовнішнім середовищем; прикладом зовнішнього середовища кінцевого автомата є множина вхідних слів, а поведедінкою - словникова функція, що реалізується автоматом, або подія, що представляється автоматом.

Крім перерахованих, в теорії автоматів є специфічні проблеми, характерні для автоматів. Так, в залежності від умов задачі поведінку автомата зручно задавати різними мовами, в зв'язку з чим, важливими задачами є вибір досить зручної адекватної мови і переклад з однієї мови на іншу. У тісному зв'язку із задачами синтезу і еквівалентних перетворень знаходиться задача мінімізації числа станів автомата, а також отримання відповідних оцінок. Спеціальний розділ теорії автоматів пов'язаний з так званими експериментами з автоматами (тобто способами отримання інформації про внутрішню структуру автоматів з їхньої поведінки). Основна задача тут полягає в тому, щоб отримати певні відомості про структуру автомата шляхом спостереження його реакції на ті чи інші зовнішні впливи. При цьому виникає велике коло задач, пов'язаних з класифікацією експериментів і з питаннями можливості розв'язання задач певними видами експериментів, а також з оцінками довжин мінімальних експериментів, достатніх для вирішення тих чи інших задач. Поняття експерименту з автоматами використовується також в задачах надійності і контролю керуючих систем, зокрема контролю автоматів. Багато з перерахованих вище задач можуть розглядатися як алгоритмічні проблеми. Для кінцевих автоматів більшість з них мають позитивне рішення.

Теорія автоматів знаходить широке застосування, як в областях математики, так і у вирішенні практичних задач. Наприклад, засобами теорії автоматів доводиться розв’язність деяких формальних числень. Застосування методів і понять теорії автоматів до вивчення формальних і – математичних дисциплін, предметом якої є розробка формального апарату для опису побудови природних і деяких штучних мов. Поняття автомата може слугувати модельним об'єктом в найрізноманітніших задачах, завдяки чому можливе застосування теорії автоматів в різних наукових і прикладних дослідженнях.

4.4.1. Еквівалентні перетворення автоматів

Два автомата називаються еквівалентними, якщо вони мають однакові вхідні і вихідні алфавіти і на однакові вхідні слова видають однакові вихідні слова.

Перехід від автомата Мілі до еквівалентного автомата Мура:

Закони функціонування автоматів Мілі і Мура відрізняються функцією виходів. Оскільки кожній парі "стан - вхідний сигнал" автомата Мілі може відповідати свій вихідний сигнал, а в автоматі Мура вихідний сигнал приписується станам, то кожній парі $(a_i;x_j)$ автомата Мілі ставиться у відповідність стан синтезованого автомата Мура b_{ij}. Таким чином, кожній клітинці таблиці переходів автомата Мілі буде відповідати новий стан автомата Мура. Крім того, оскільки перший вихідний сигнал автомат видасть тоді, коли з початкового стану під впливом першого вхідного сигналу він перейде в якийсь стан, який і визначить перший вихідний сигнал, то необхідно для автомата Мура ввести також початковий стан b_0, якому може бути приписаний будь-який допустимий вихідний сигнал. Оскільки функції переходів у автоматів Мілі і Мура однакові, кожному стану автомата Мілі ставиться у відповідність клас ізоморфних станів автомата Мура.

Таким чином, існує стандартний прийом, за допомогою якого можна перетворити автомат Мілі в еквівалентний йому автомат Мура. Причому, якщо в автоматі Мілі було n внутрішніх станів і кількість вхідних сигналів m, то в отриманому автоматі Мура буде $(n * m) +1$ станів.

Перехід від автомата Мура до еквівалентного автомата Мілі:

З аналізу переходів автоматів Мілі і Мура слідує, що для переходу від автомата Мілі до автомата Мура необхідно як би віднести кожен вихідний сигнал до попереднього стану і вхідного сигналу, який перевів автомат Мура в даний стан. Іншими словами, на графі автомата вихідні сигнали, раніше приписані вершинам, необхідно віднести до відповідних дуг, що заходять в цю вершину.

4.4.2. Мінімізація автоматів

Мінімальний автомат – це автомат, який має найменшу можливу кількість станів, і який реалізує задану функцію виходів.

Два стани автомата називаються *1-евівалентними*, якщо, перебуваючи в будь-якому із цих станів, автомат на один і той же вхідний сигнал (вхідне слово довжиною в 1) видає один і той же вихідний сигнал. Два стани автомата називаються *К-еквівалентними*, якщо, починаючи з будь-якого із цих станів, автомат на будь-які однакові слова довжини *К* видає однакові вихідні слова (також довжини *К*). Якщо два стани *К*-еквіваленті для будь-яких *К*, то їх називають (просто) еквівалентними. Множина попарно еквівалентних станів формує клас еквівалентності.

Зміст мінімізації полягає у виявленні класів еквівалентності і заміні кожного класу одним станом. Процедура мінімізації полягає у наступному:

1) за таблицею виходів автомата Мілі (або за першою строфою відміченої таблиці переходів автомата Мура) знаходяться стани, які мають однакові стовпчики відмічені однаковими вихідними сигналами – це 1-еквівалентні стани.

2) далі використовується таблиця переходів. Всі стани, що входять до 1-еквівалентного класу, і, що під дією першого сигналу перейшли в стани, які у свою чергу належать до 1-еквівалентного класу, формують 2-еквівалентний клас.

3) процедура розподілу на класи еквівалентності продовжується до того часу, доки при черговому кроці К-еквівалентні класи не співпадуть із (К-1) - еквівалентними, тобто поки не отримаємо еквівалентні класи.

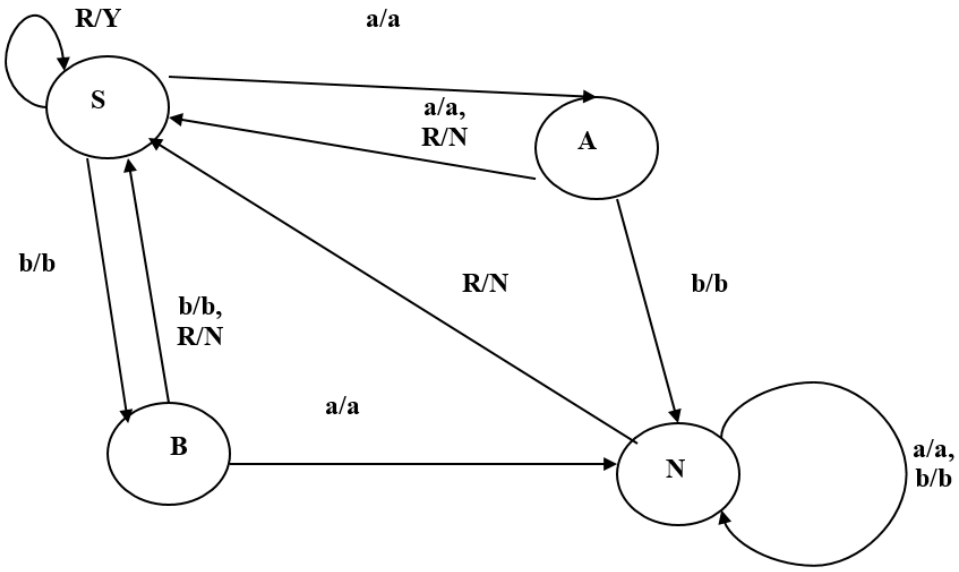
4) всі стани, що входять в один клас еквівалентності, замінюються одним станом.

4.4.3. Розпізнавальні автомати

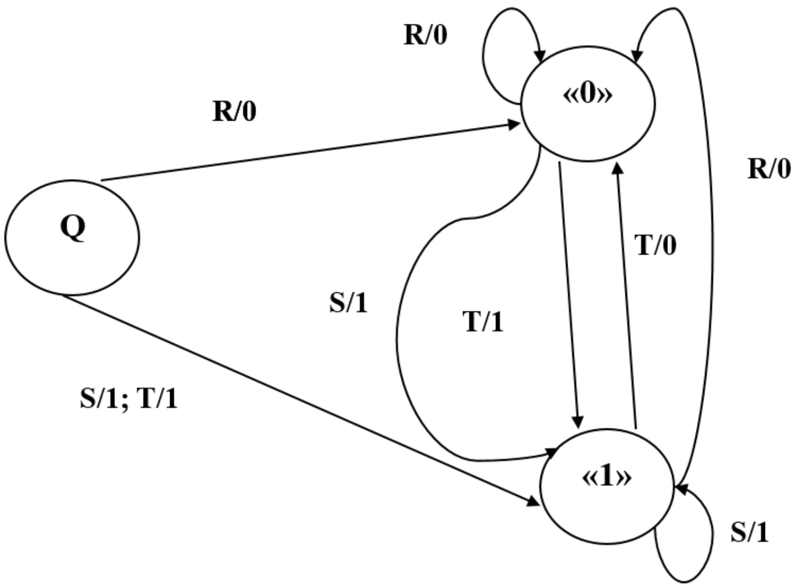
Розпізнавальний автомат – це автомат Мура, в якому фіксуються початковий стан і підмножина станів F×Q, яка називається множиною кінцевих станів. Кажуть, що автомат допускає (приймає, розпізнає, представляє) дане слово, якщо реакцією на це слово може бути перехід автомата в один із кінцевих станів.

Приклади. Побудувати (синтезувати) автомат за змістовим описом.

1. Автомат (розпізнавальний) з початковим станом S(start) допускає слова, у яких є лише парні входження літер a і b, наприклад, a a, a a a a a a, b b a a і т. п. Ознакою кінця слова є сигнал R. Тобто у випадку вхідного слова aaaabbaaR видається вихідне слово aaaabbaaY (тобто так), а у випадку aaaabaaR – aaaabaaN (тобто ні).



2. На вхід автомата можуть поступати сигнали R, S і T. На вхідний сигнал R автомат видає вихідний сигнал 0, на S – вихідний сигнал 1, і на T – вихідний сигнал, протилежний попередньому вихідному сигналу. Для визначеності вважаємо, що в початковому стані автомат пам’ятає «попередній» вихідний сигнал 0.



3. Одним із найбільш використовуваних на практиці типів розпізнавальних автоматів є частковий недетермінований автомат. Недетермінізм проявляється в тому, що із одного стану за одним із тим самим вхідним сигналом можливі переходи в різні стани, тобто функція переходів заміняється відношенням переходів. Недетермінований автомат приймає, наприклад, слова ab, aa, bb, bba і т.п. Тут початковий стан A, кінцевий – F. Таблиця переходів даного автомата буде мати вигляд:

	A	**B**	**C**	**F**
a	B,C	-	F	-
b	B	C,F	-	-

Порожні клітинки свідчать про те, що автомат частковий, а наявність одразу декількох літер в одній клітинці – про те, що автомат недетермінований.

Для таких автоматів зазвичай надають перевагу не табличному і графомову представленню, а запису у вигляді, так званих продукцій, чи граматичних правил, що представлені у двох стовпчиках відповідно:

A → aB A:: = aB / bB / aC

A → bb B:: = bC / b

A → aC C:: = a

B → bC

B→b

C→a

Граматики такого роду називають регулярними або автоматними. Автомати такого типу будуть розглядатися у зв’язку із вивченням теорії формальних мов і граматик.

Питання для самоконтролю до розділу 4

1. Абстрактні скінченні автомати.
2. Функції переходу та виходу.
3. Способи завдання автоматів.
4. Автомати Мілі.
5. Автомати Мура.
6. Особливості теорії автоматів: сінтез, аналіз; різні види автоматів.
7. Еквівалентні перетворення автоматів.
8. Особливості перетворювання автомата Мілі в автомат Мура.
9. Мінімізація автоматів. Еквівалентні стани.
10. Розпізнавальні автомати.

5. МЕРЕЖІ ПЕТРІ

5.1. Основні положення. Дозвіл і запуск переходів

Мережі Петрі - математичний апарат для моделювання динамічних дискретних систем. Вперше описані Карлом Петрі в 1962 році.

Мережа Петрі це дводольний орієнтований граф, що складається з вершин двох типів - позицій і переходів, з'єднаних між собою дугами, вершини одного типу не можуть бути з'єднані безпосередньо. У позиціях можуть розміщуватися мітки (маркери), здатні переміщатися по мережі.

Подією називають спрацьовування переходу, при якому мітки з вхідних позицій цього переходу переміщаються в вихідні позиції. Події відбуваються миттєво, різночасно при виконанні деяких умов.

Мережі Петрі - інструмент дослідження систем. Розвиток теорії мереж Петрі проводилося за двома напрямками. Формальна теорія мереж Петрі займається розробкою основних засобів, методів і понять, необхідних для застосування мереж Петрі. Прикладна теорія мереж Петрі пов'язана головним чином із застосуванням мереж Петрі до моделювання систем, їх аналізу.

Моделювання в мережах Петрі здійснюється на подієвому рівні. Визначаються, які дії відбуваються в системі, які стани передували цим діям і які стани прийме система після виконання дії. Виконання подієвої моделі в мережах Петрі описує поведінку системи.

Мережа Петрі, як вже було сказано вище, являє собою деякий різновид орієнтованого графа із заданим початковим станом, який називається початковим маркуванням (або розміткою) M_0. Граф *n* мережі Петрі є орієнтованим і включає вузли (або вершини) двох типів, звані позиціями (або місцями) і переходами, дуги в якому ведуть або з позиції в перехід, або з переходу в позицію. У графічному поданні позиції зображуються кружками, а переходи жирними рисками або прямокутниками. Дуги позначаються відповідними вагами (цілими позитивними числами), і дугу з вагою *k* можна вважати еквівалентної *k* паралельним дуг. Вказівка одиничний ваги зазвичай опускається. Маркування (стан) приписує кожній позиції ціле невід'ємне число. Якщо маркування приписує позиції p ціле невід'ємне число *k*, то говорять, що *p* марковано *k* фішками; в графічному зображенні всередині позиції-кружка в цьому випадку поміщають *k* чорних точок (фішок). Маркування позначається вектором *M* довжиною *m*, де *m*- загальне число позицій. Компонента вектора *M* з номером *p*, що позначається *M (p),* дорівнює числу фішок в позиції *p*.

У задачах моделювання, де застосовуються поняття умов (станів) і подій, позиції відповідають умовам, а переходи - подіям. Кожен перехід (подія) пов'язаний з певним числом вхідних і вихідних позицій - аналогів відповідно передумов та постумов цієї події. Наявність фішки в деякій позиції інтерпретується як істинність умови, що відповідає даній позиції. В іншому трактуванні позиція позначається *k* фішками, з тим, щоб вказати на наявність *k* елементів даних або відповідної кількості ресурсів. У табл.5.1 представлено кілька типових варіантів інтерпретації переходів і відповідних їм вхідних і вихідних позицій.

Таблиця 5.1.

Вхідні позиції	**Перехід**	**Вихідні позиції**
Передумова	Подія	Постумова
Вхідні дані	Обчислювальна операція	Вихідні дані
Вхідні сигнали	Процесор сигналу	Входні сигнали
Затребувані ресурси	Задача	Вивільнені ресурси
Умова	Логичні вислови (клозы)	Висновки
Буфери	Процесор	Буфери

Формальне визначення мережі Петри (P_n)

Мережа Петрі – це п'ятірка $P_n = (P, T, F, W, M_0)$, де

$P = (p_1, p_2, \ldots, p_m)$ – кінцева множина позицій;

$T = (t_1, t_2, \ldots, t_n)$ – куінцева множина переходів;

$F \subseteq (p \times t) \cup (t \times p)$ – множина дуг;

$W : f \to (1, 2, 3, \ldots)$ – вагова функція;

$M_0 : p \to (0, 1, 2, 3,..)$ – початкова маркіровка;

$P \cap t = \varnothing$.

Структура мережі Петрі $N=(P,T,F,W)$ без заданого початкового маркування позначається однією літерою N, с заданим – двома (N, M_0).

При моделюванні динаміки системи стан, або маркування мережі Петрі, змінюється відповідно до правила (запуску) переходу:

1. Перехід дозволено, якщо все вхідні позиції р переходу помічені не менше ніж *w (p, t)* фішками, де *w (p, t)* - вага дуги, що веде з *p* в *t*;

2. Запуск дозволеного переходу носить випадковий характер (в залежності від настання або ненастання відповідної події);

3. Запущений перехід *t* вилучає *w(p, t)* фішок з кожної своєї вхідної позиції і додає *w (t, p)* фішок в кожну свою вихідну позицію; тут *w (t, p)* - вага дуги, що веде з *t* в *p*.

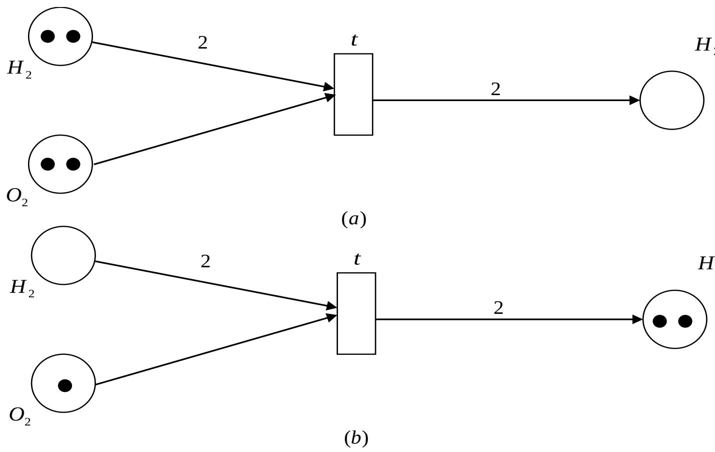


Рис. 5.1. Пояснення до правила запуску переходу: (a) маркування перед спрацьовуванням дозволеного переходу t; (b) маркування після спрацьовування цього переходу, після чого він стає забороненим переходом.

Перехід, який не має жодної вхідної позиції, називається витоком, а перехід, який не має жодної вихідної позиції – стоком. Відзначимо, що перехід-витік є, безумовно, дозволеним переходом, а запуск переходу-стоку призводить до вилучення фішок, що не породжуючи нових. (рис.5.1)

Пара, що складається з позиції *p* і переходу *t*, називається петлею, якщо *p* служить одночасно вхідний і вихідний позицією переходу *t*. Мережа Петрі називається однорідною, якщо такі пари в ній відсутні. Мережа Петрі називається ординарною, або простою, якщо вага будь-якої її дуги дорівнює 1.

5.2. Найпростіші приклади модельованих об'єктів

5.2.1. Кінцеві автомати

Кінцевий автомат або відповідну йому діаграму станів можна адекватно представити деяким подклассом мереж Петрі. Як приклад кінцевого автомата розглянемо торговий автомат, який приймає монети гідністю в 5 і 10 копійок і продає цукерки вартістю 15 і 20 копійок. Для спрощення будемо вважати, що ємність монето-приймача обмежена 20 копійками. Тоді діаграма станів автомата може бути представлена мережею Петрі, показаної на рис.5.2.

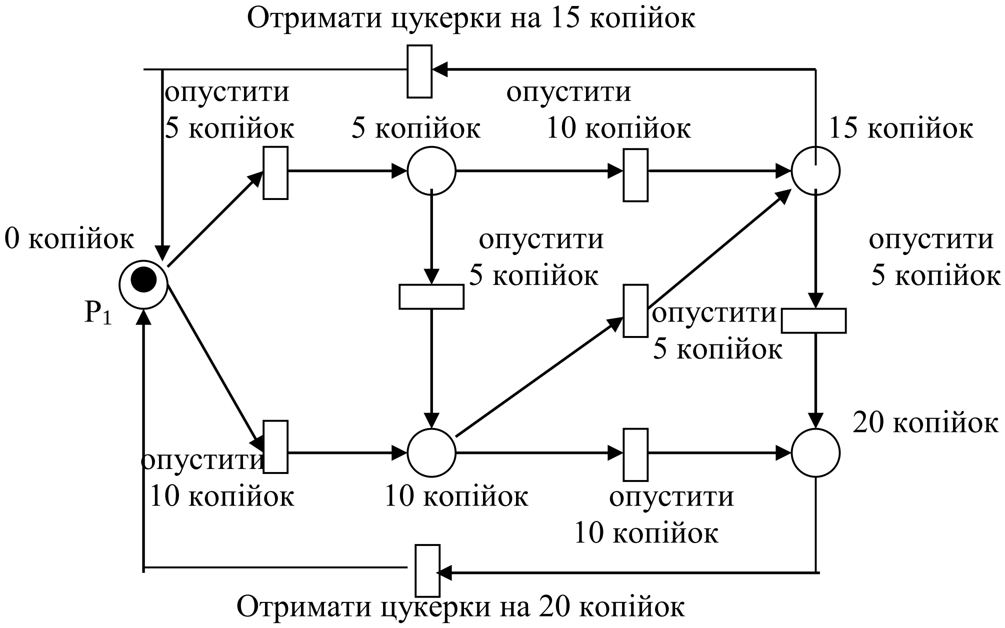


Рис. 5.2. Мережа Петрі (автоматна мережа), що моделює діаграму станів торгового автомата. Переходи, що відповідають поверненню монет, пропущени для спрощення.

На малюнку п'яти станам відповідають п'ять позицій, позначених мітками 0, 5, 10, 15, 20 копійок, а переходами з одного стану в інший - переходи, позначені вхідними умовами типу «опустити 5 копійок». Початковий стан відзначено приміщенням фішки в позицію P_1, позначену в даному прикладі міткою «0 копійок». Зауважимо, що в цій мережі кожен перехід має тільки одну вхідну і тільки одну виходну дуги. Підклас мереж Петрі, що володіє такою властивістю, називається автоматним. Будь-який кінцевий автомат (або його діаграму станів) можна представити деякою автоматної мережею.

Фрагмент з позицією P_1, яка передує двом (або більше) вихідним переходам t_1 і t_2 (рис.5.3), називається, в залежності від концепта, конфліктом або вузлом прийняття рішення.

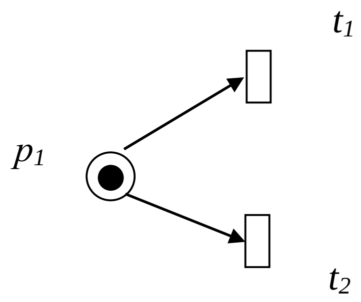


Рис. 5.3. Фрагмент мережі Петрі, званий конфліктом, вибором або вузлом прийняття рішення. Ця структура має властивість недетермінованости.

Автоматні мережі дозволяють відображати прийняття рішень, але вони не придатні для моделювання синхронізації паралельних процесів.

5.2.2. Паралельні процеси

За допомогою мереж Петрі зручно відображати паралельні процеси, або паралелізм. Наприклад, в мережі на рис.5.4 паралельні процеси, яким відповідають переходи t_2 і t_3, починаються при запуску переходу t_1 і завершуються при запуску переходу t_4.

У загальному випадку два переходи вважаються паралельними, якщо між ними відсутній причинно-наслідковий зв'язок, тобто один перехід може спрацювати раніше іншого, після нього або одночасно з ним, як, наприклад, в разі переходів t_2 і t_3 на рис.5.4.

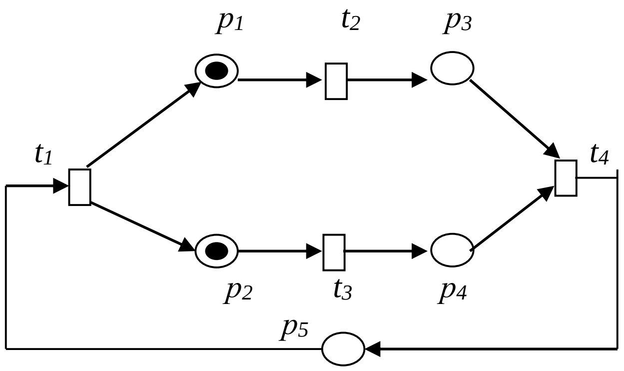


Рис. 5.4. Мережа Петрі (маркований граф), що моделює детерміновані паралельні процеси.

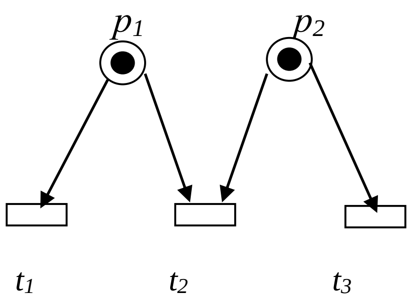
Паралелізм можна розглядати як бінарне відношення (позначається символом х і діє на множині $A = \{e_1, e_2, \dots \}$), яке є рефлексивним, тобто e_i х e_i, і симетричним, тобто з e_1 х e_2 слідує e_2 х e_1, але не володіє транзитивністю, тобто з e_1 х e_2 і e_2 х e_3 не обов'язково слідує e_1 х e_3. Наприклад, можна вести машину (подія e_1) або йти пішки (подія e_3) і при цьому співати (подія e_2), проте не можливо одночасно (паралельно) вести машину і йти пішки.

Зауважимо, що кожна позиція в мережі, показаної на рис.5.4 має тільки одну вхідну і одну вихідну дуги. Підклас мереж Петрі, що володіють такою властивістю, називається маркованими графами. Вони дозволяють представляти паралельні процеси, але не придатні для моделювання прийняття рішень (конфліктів).

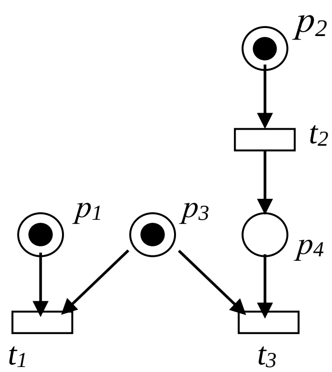
Події e_1 і e_2 є конфліктуючими, якщо вони обидва можливі, але не можуть наступити одночасно. Події e_1 і e_2 паралельні, якщо вони можуть наступити в будь-якій послідовності, не приводячи до конфліктів. Поєднання конфліктної ситуації з паралелізмом називається змішанням.

На рис.5.5 показано два типи змішування.

При симетричному змішуванні (рис.5.5.а) події t_1 і t_2 паралельні, і обидва знаходяться в конфлікті з подією t_3. При асиметричному змішуванні (рис.5.5.б) подія t_1 паралельно t_2, і вступить в конфлікт з t_3, якщо t_2 спрацює раніше.



а)



б)

Рис.5.5. Два типу змішування: а) симетричне змішування, б) асиметричне змішування.

5.2.3. Обчислення, що управляються потоками даних

Мережі Петрі придатні для моделювання потоку управління, а й потоку даних. На рис.5.6 показана мережа Петрі, що моделює обчислювальну операцію, керовану потоком даних. У ЕОМ, керованої потоком даних, виконання операції ініціюється надходженням її операндів і може здійснюватися паралельно. У мережі Петрі, що моделює таку організацію обчислювального процесу, фішки показують значення поточних даних, а також наявність даних. У мережі на рис.5.6 операції відповідні переходам t_1 і t_2, можуть виконуватися паралельно, а їх результат, тобто *(а+в)* і *(а–в),* віддається в відповідні вихідні позиції.

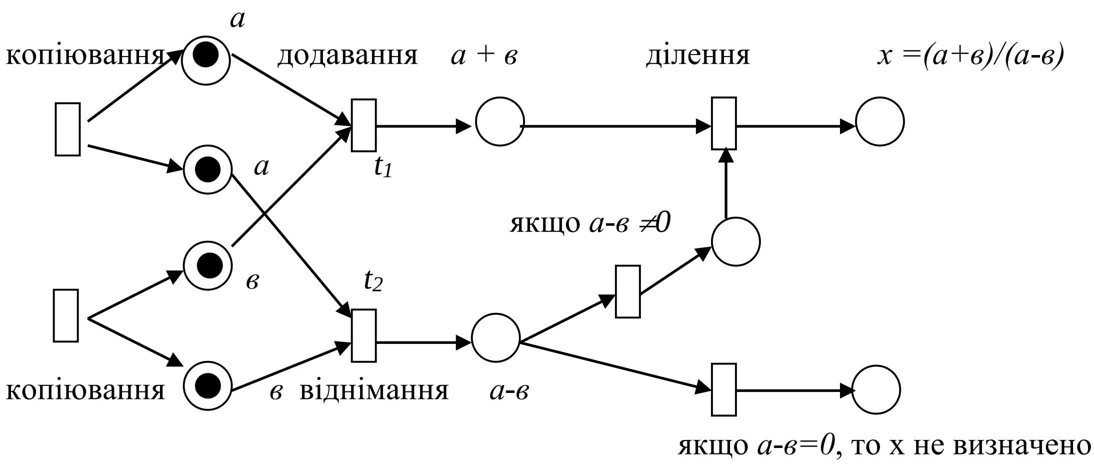


Рис.5.6. Мережа Петрі, що моделює обчислювальну операцію
х = (а+в) / (а–в), керовану потоком даних.

5.2.4. Система виробники-споживачі з пріоритетами

Мережа, показана на рис. 5.7, моделює систему виробники-споживачі з пріоритетами, в якій споживач *а* має вищий пріоритетам в порівнянні зі споживачем *в* в тому сенсі, що поки буфер *а* містить деякі об'єкти (фішки), споживач *а* може їх безперешкодно отримувати .

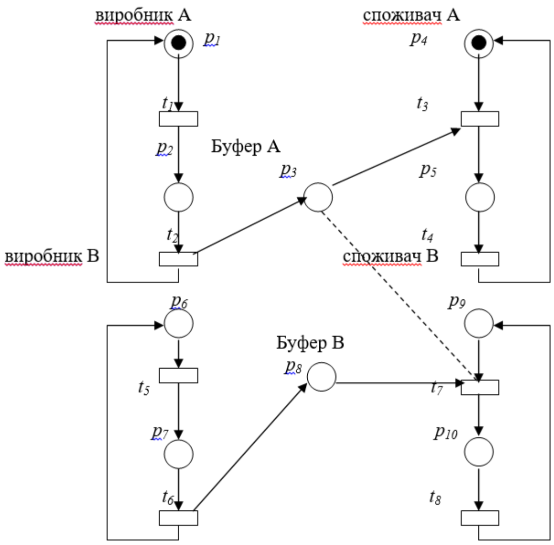


Рис. 5.7. Розширена мережа Петрі, що моделює систему виробники-споживачі з пріоритетами.

5.3. Поведінкові властивості

Великою перевагою мереж Петрі є можливість аналізу з їх допомогою багатьох властивостей і завдань, пов'язаних з паралельними системами. Моделі на основі мереж Петрі дозволяють досліджувати два види властивостей: ті, що визначаються початковим маркуванням, і ті, що не залежать від неї. Властивості, що відносяться до першої групи, називаються поведінковими, а властивості другої групи - структурними.

- *Досяжність*

Досяжність є фундаментальним поняттям, необхідних для вивчення динамічних властивостей будь-якої системи. Запуск переходу, що дозволено, призводить до зміни розподілу фішок в мережі (зміні маркування) за правилом переходу. Послідовності запусків відповідає послідовність маркувань. Маркування M_n досяжно від маркування M_0, якщо існує послідовність запусків, що призводять від M_0 до M_n. Послідовність запусків (подій) позначається як $\sigma = м_0 t_1\ м_1 t_2\ м_2 \ldots t_n м_n$, або просто $\sigma = t_1\ t_2 \ldots t_n$,. При цьому говорять, що M_n досяжно від M_0 в результаті послідовності σ, і цей факт записується у вигляді наступного виразу: $M_0[\sigma > M_n]$. Множина всіх маркувань, досяжних в мережі (n, M_0) від M_0, позначається через $r\ (n, M_0)$, або просто $r(M_0)$.

Множина всіх послідовностей запусків в мережі (n, M_0), що починаються з маркування M_0, позначається через $l\ (n, M_0)$, або просто $l\ (M_0)$.

Таким чином, проблема досяжності в мережах Петрі полягає в тому, щоб для заданого маркування M_n в мережі (n, M_0) встановити приналежність $M_n \in r(M_0)$. У деяких додатках може виникнути необхідність дослідити маркування тільки деякої підмножини позицій мережі. При цьому виникає проблема досяжності подмаркіровкі, яка полягає у встановленні приналежності $M'_n \in r(M_0)$, де M'_n – будь-яке маркування, що збігається із заданим маркуванням M_n з точністю до маркування позицій, що входять в задану підмножину. Проблема досяжності вирішувана, але в загальному випадку витрати пам'яті (і часу) на її рішення зростають принаймні експоненціально. Однак проблема рівності є нерозв'язною, тобто, не існує алгоритму, який для будь-яких мереж Петрі n і n' може встановити виконання рівності $l(n, M_0) = l(n', M'_0)$.

- ***Обмеженість***

Сеть Петри называется *k-ограниченной*, или просто *ограниченной*, если для любой маркировки, достижимой от маркировки $м_0$, количество фишек в любой позиции не превышает некоторого конечного числа k,т.е. $M(p) \leq k$ для любого p и любой маркировки $м \in r\ (м_0)$. Сеть петри $(\ n,\ м_0\ )$, называется *безопасной*, если она 1-ограниченна. Позиции сети петри часто моделируют буферы или регистры, служащие для хранения промежуточных результатов. Проверка ограниченности или безопасности сети гарантирует отсутствие переполнения буферов или регистров при любой фактической последовательности запусков переходов.

Мережа Петрі називається k-обмеженою, або просто *обмеженою*, якщо для будь-якого маркування, що є досяжним від маркування M_0, кількість фішок в будь-якій позиції не перевищує деякого кінцевого числа k, тобто, $M(p) \leq k$ для будь-якого p і будь-яке маркування $M \in r\ (M_0\ )$. Мережа Петрі (n, M_0), називається *безпечною*, якщо вона *1-обмежена*. Позиціями мережі

Петрі часто моделюють буфери або регістри, що служать для зберігання проміжних результатів. Перевірка обмеженості або безпеки мережі гарантує відсутність переповнення буферів або регістрів при будь-якій фактичної послідовності запусків переходів.

- ***Активність***

Поняття *активності* тісно пов'язане з відсутністю можливості взаємного блокування (тупикових ситуацій) в операційній системі. Мережа Петрі *активна*, якщо, незалежно від досягнутого від M_0 маркування, для будь-якого переходу існує послідовність подальших запусків, яка веде до його запуску. Це означає, що для активної мережі Петрі при будь-якій послідовності запусків повністю виключена можливість взаємного блокування.

- ***Зворотність і базовий стан***

Мережа Петрі зворотна, якщо для будь-якої маркування M з $r(M_0)$ маркування M_0 досяжно від M. Іншими словами, в зворотній мережі завжди можна повернутися до початкового маркування (стану). У багатьох задачах не обов'язково повертатися в початковий стан, але досить мати можливість повернутися в якийсь базовий стан. Тому умову зворотності можна послабити і визначити поняття базового стану. Маркування M' називається базовим станом, якщо вона досяжна від будь-якого маркування M из $r(M_0)$.

- ***Стійкість***

Мережу Петрі (n, M_0) називається стійкою, якщо для двох будь-яких дозволених переходів запуск одного з них не призводить до дозволу спрацювання іншого. В стійкої мережі будь-який перехід, що став дозволеним, зберігає цей стан до тих пір, поки не спрацює. Поняття стійкості виявляється корисним при вивченні паралельних схем обчислень і асинхронних схем. Стійкість тісно пов'язана з поняттям безконфліктних мереж, а безпечно стійка мережа може бути перетворена в маркований граф за допомогою копіювання деяких переходів і позицій. Зауважимо, що хоча всі марковані графи є стійкими, зворотне твердження в загальному випадку не справедливо.

5.4. Методи аналізу мереж Петрі

Методи аналізу мереж Петрі можна розбити на наступні три групи:

1) методи на основі дерева покриваемості (дерева досяжності);

2) методи алгебри на основі матричних рівнянь;

3) методи перетворення або декомпозиції.

Методи першої групи, по суті справи, зводяться до перерахування всіх досяжних маркувань або відповідних їм маркувань, що покриваються, і, взагалі кажучи, можуть бути застосовані до будь-яких класів мереж, однак на практиці їх використання обмежене «невеликими» мережами через різке зростання складності і потужності простору станів при збільшенні розмірів мережі. У той же час матричні рівняння і методи перетворення мереж, будучи могутнім засобом аналізу, виявляються придатними тільки для окремих підкласів мереж Петрі або лише в окремих випадках.

5.4.1. Дерево покриваемості

Для даної мережі Петрі з початковим маркуванням M_0 число нових можливих маркувань дорівнює числу дозволених переходів. Від кожного наступного маркування можна знову перейти до нових маркувань. Такий процес зміни маркувань можна представити у вигляді дерева. Його вузли відповідають маркуванням, що породжується маркуванням M_0 (коренем) і наступними вузлами, а дуги - спрацьовуванню переходів, що призводять до зміни маркувань.

Однак, в разі необмеженої мережі таке подання у вигляді дерева буде розростатися до нескінченності. Для того, щоб уникнути цього, введемо спеціальний елемент ω, що володіє наступними властивостями: для будь-якого цілого n має місце $\omega > n$, $\omega \pm n = \omega$ і $\omega \geq \omega$, тобто це нескінченно велике число.

Розглянемо приклад по наступній мережі Петрі (рис.5.8):

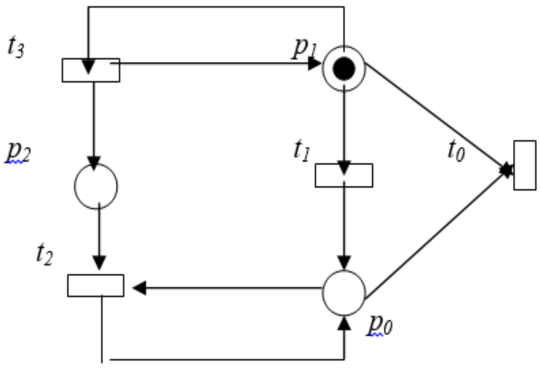
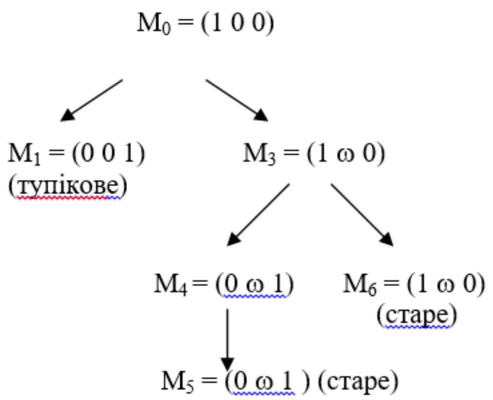


Рис.5.8. Перехід t_0 пасивний, t_1 і t_2 – активні.

При начальному маркуванні M_0 = (1 0 0) маємо два дозволених переходи: t_1 і t_3. Запуск перехода t_1 змінює M_0 на M_1 = (0 0 1), що представляє собою «тупіковий» вузол, тому що при маркуванні M_1 не існує дозволених переходів. Запуск перехода t_3 при маркуванні M_0 призводить к появі маркування $M_3{}'$ = (1 1 0), що покриває M_0 = (1 0 0). Так як новим стає маркування $M_3{}'$ = (1 ω 0), при якому переходи t_1 і t_3 знову виявляються дозволеними. Запуск переходу t_1 змінює M_3 на M_4 = =(0 ω 1), при цьому може спрацювати переход t_2 призводя до «старого» маркування M_5= M_4. Запуск переходу t_3 при маркуванні M_3 веде до «старого» вузла M_6= M_3. Таким чином, дерево покриваемості, що показане на рис.5.9, побудовано.



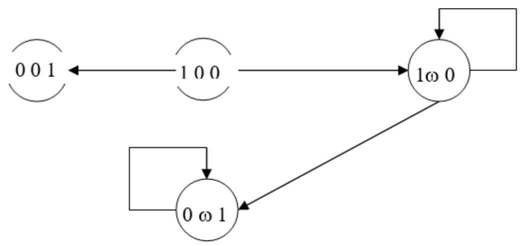


Рис.5.9. Дерево покриваемості та граф досяжності мережі з рис.5.8.

За допомогою дерева покриваемості *T* можна, наприклад, досліджувати такі властивості мережі Петрі *(n, M_0)*:

1. Мережа *(n, M_0)* обмежена і, отже, множина *r (M_0)* звичайно тоді і тільки тоді, коли ω не входить не в одну мітку вузла *T*;

2. Мережа *(n, M_0)* безпечна тоді і тільки тоді, коли в мітки вузлів *T* входять тільки *1* і *2*;

3. Перехід *t* пасивний тоді і тільки тоді, коли він не є міткою дуги дерева Т;

4. Якщо маркування М досяжно від маркування M_0, то існує вузол з міткою *M'*, такий що $M_0 \le M'$.

У разі обмеженою мережі Петрі дерево покриваемості називається деревом досяжності, оскільки воно містить всі досяжні маркування мережі.

Недолік описаного підходу полягає в тому, що це метод повного перебору. Однак в загальному випадку проблеми активності і досяжності не можуть бути вирішені за допомогою одного тільки методу покриваемості, оскільки частина інформації втрачається при підстановці символу ω (який може відповідати парним або непарним, зростаючим або спадним числам і т.п.).

5.4.2. Матриця інцидентності та рівняння стану

Динаміку багатьох систем, що розглядаються в техніці, можна описати диференціальними або алгебраїчними рівняннями. Розглянемо можливість отримання рівнянь, що дозволяють повністю описати і дослідити динаміку мереж Петрі. Нижче представлені матричні рівняння, що визначають динаміку паралельних процесів, що моделюються мережами Петрі. Однак ці рівняння не завжди можна розв'язати, що пов'язано, з одного боку, з недетермінірованностью моделей на основі мереж Петрі, а з іншого боку, - обмеженням, що вимагають рішення в вигляді невід'ємних цілих чисел.

Матриця інцидентності. Для мережі Петрі *n*, що має *n* переходів і *m* позицій, матрицею інцидентності $a=[a_{ij}]$ називається $(n * m)$ - матриця, елементами якої є цілі числа, що визначаються виразом

$A_{ij} = a_{ij}^{+} - a_{ij}^{-}$,

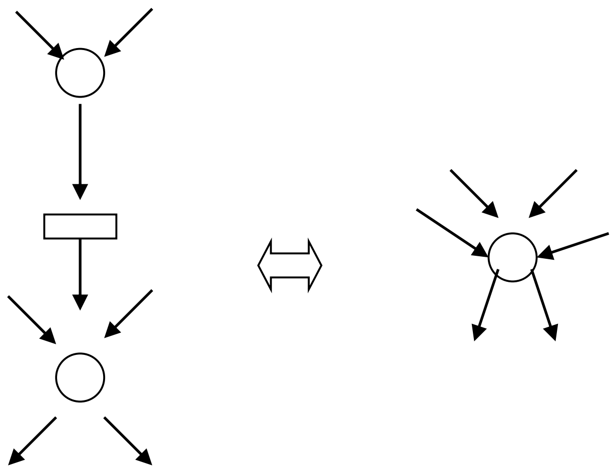
де $a_{ij}^{+} = \omega(i,j)$, вага дуги, що веде від переходу *i* до вихідний позиції *j*,

а $a_{ij}^{-} = \omega(i,j)$, вага дуги, що веде від переходу *i* до вхідних позиції *j*.

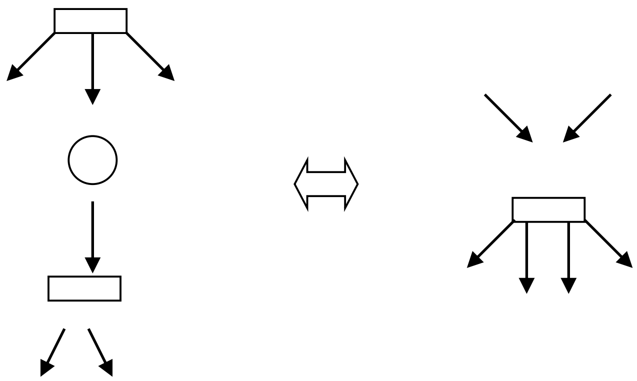
Рівняння стану. У матричних рівняннях маркування $м_k$ записується у вигляді $(m * 1)$ - вектора-стовпця. Елемент *j* вектора $м_k$ дорівнює числу фішок в позиції j відразу після запуску k-го переходу в деякій послідовності i запусків. K-й запуск, або вектор управління u_k, являє собою $(n * 1)$ - вектор - стовпець, всі елементи якого, крім елемента j, рівні 0.

5.4.3. Аналіз за допомогою простих правил перетворювання мереж

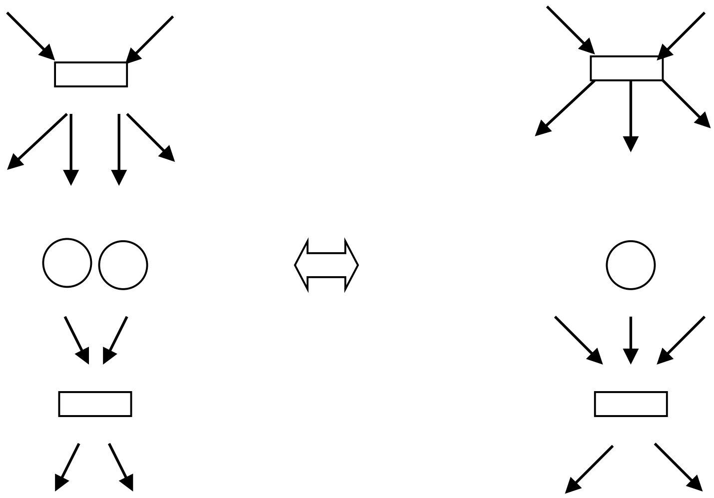
Для спрощення аналізу складної системи часто призводять до більш простої моделі, що зберігає властивості досліджуваної системи. І навпаки, для задач синтезу можна використовувати методи послідовного перетворення абстрактної моделі в більш детальну. Існує цілий ряд перетворень мереж Петрі.



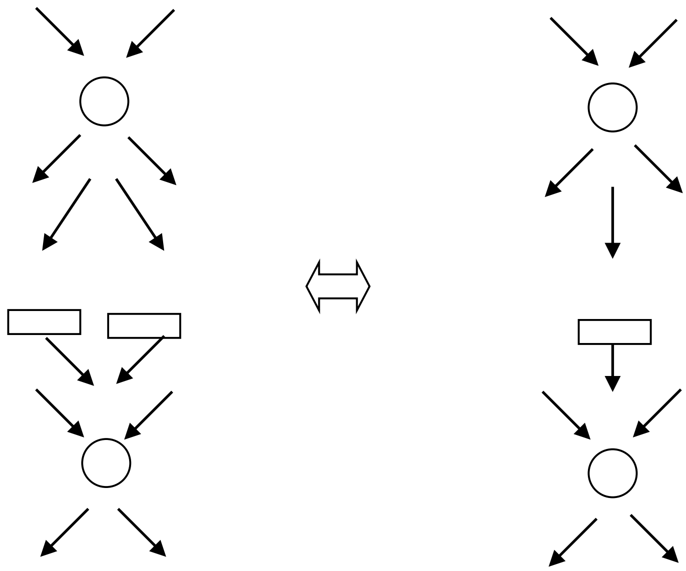
а) операція злиття послідовно з'єднаних позицій;



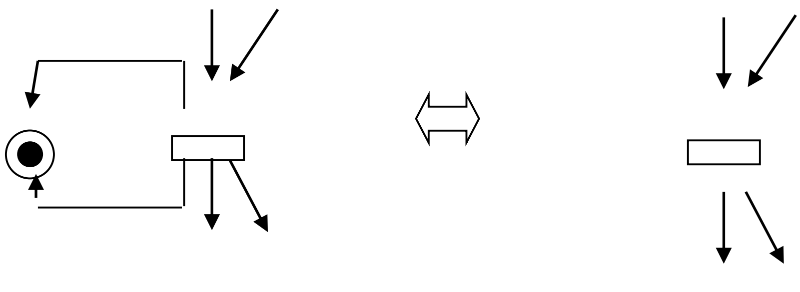
б) операція злиття послідовно з'єднаних ппереходів;



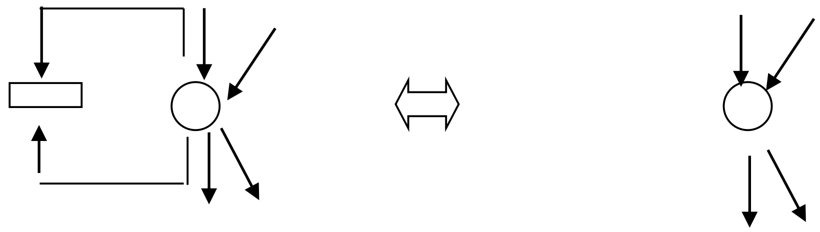
в) операція злиття паралельних позицій;



г) операція злиття паралельних переходів;



д) операція виключення позицій з петлями;



е) операція виключення переходів з петлями

Рис. 5.10. Шість операцій перетворення, що зберігають властивості активності, безпеки і обмеженості

5.5. Структурні властивості

Структурними властивостями мереж Петрі називаються властивості, що залежать від їх типології. Ці властивості не залежать від початкового маркування $м_0$ в тому сенсі, що вони зберігаються при будь-якому початковому маркуванні або, іншими словами, пов'язані з існуванням певних послідовностей запуску (спрацьовування) переходів, ініційованих з деякого початкового маркування. Тому ці властивості можна, як правило, охарактеризувати за допомогою матриці інцидентності a й асоційованих з нею однорідних рівнянь або нерівностей.

С*труктурна активність*. Мережа Петрі n називається структурно активної, якщо для неї існує деяка активна початкова розмітка.

Керованість. Якщо мережа Петрі n з m позиціями керована, то $Rank\ a=m$.

Структурна обмеженість. Мережа Петрі n називається структурно обмеженою, якщо вона є обмеженою для будь-якої кінцевої початкової розмітки $м_0$.

Консервативність. Мережа Петрі n називається консервативної (частково консервативної), якщо для кожної (деякою) позиції p існує позитивне ціле число $y\ (p)$, таке що для будь-якої маркування $м \in r\ (м_0)$ і будь-якій фіксованій початковій маркування $м_0$ зважена сума фішок $M^m y = м^m{}_0 y = const$.

Консистентність. Мережа Петрі *n* називається консистентною (частково консистентною), якщо існує маркування $м_0$ і послідовність запуску переходів υ з $м_0$ назад в $м_0$, таке що будь-який (деякий) перехід хоча б один раз спрацьовує в υ.

5.6. Застосування мереж Петрі

Мережі Петрі широко застосовуються в системах масового обслуговування та імітаційному моделюванні. Система масового обслуговування – об'єкт (підприємство, організація та ін.), діяльність якого пов'язана з багаторазовою реалізацією виконання якихось однотипних завдань і операцій.

Імітаційне моделювання – це окремий випадок математичного моделювання. Існує клас об'єктів, для яких з різних причин не розроблені аналітичні моделі, або не розроблені методи рішення отриманої моделі. У цьому випадку математична модель замінюється імітатором або імітаційною моделлю.

До імітаційного моделювання вдаються, коли:

• Дорого або неможливо експериментувати на реальному об'єкті;

• Неможливо побудувати аналітичну модель: в системі є час, причинні зв'язки, наслідок, нелінійності, стохастичні (випадкові) змінні;

• Необхідно зімітувати поведінку системи в часі.

Мета імітаційного моделювання полягає у відтворенні поведінки досліджуваної системи на основі результатів аналізу найбільш суттєвих взаємозв'язків між її елементами або іншими словами. Імітаційну модель можна розглядати як множину правил (диференціальних рівнянь, карт станів, автоматів, мереж і т.п.), які визначають, в який стан система перейде в майбутньому з заданого поточного стану.

Аналіз результатів виконання може сказати про те, в яких станах перебувала або не пробувала система, які стани в принципі не досяжні. Однак, такий аналіз не дає числових характеристик, що визначають стан системи.

Підсумуємо, що основними властивостями мережі Петрі є:

• Обмеженість - число міток в будь-якій позиції мережі не може перевищити деякого значення k.

• Безпека - окремий випадок обмеженості, $k = 1$.

• Збереженість - сталість завантаження ресурсів, $\sum A_i N_i$ постійна. Де n_i - число маркерів в $_i$-тій позиції, a_i - ваговий коефіцієнт.

• Досяжність - можливість переходу мережі з одного заданого стану (що характеризується розподілом міток) в інше.

• Жвавість - можливістю спрацьовування будь-якого переходу при функціонуванні об'єкта, що моделюється.

В основі дослідження перерахованих властивостей лежить аналіз досяжності.

Розвиток теорії мереж Петрі призвело до появи, так званих, "кольорових" мереж Петрі. Поняття кольоровості в них тісно пов'язано з поняттями змінних, типів даних, умов та інших конструкцій, більш наближених до мов програмування. Незважаючи на деяку схожість між кольоровими мережами Петрі та програмами, вони ще не застосовувалися в якості мови програмування.

Не дивлячись на описані вище переваги мереж Петрі, незручності застосування мереж Петрі в якості мови програмування укладені в процесі їх виконання в обчислювальній системі. У мережах Петрі немає строго поняття процесу, який можна було б виконувати на зазначеному процесорі. Немає також однозначної послідовності виконання мережі Петрі, так як вихідна теорія представляє нам мову для опису паралельних процесів.

Питання для самоконтролю до розділу 5

1. Визначення мережі Петрі.
2. Дозвіл і запуск переходів.
3. Моделювання конфлікту та паралельних процесів.
4. Поведінкові властивості.
5. Структурні властивості.
6. Методи аналізу мереж Петрі.
7. Дерева покриваемості.

8. Матричний метод аналізу.
9. Прості перетворювання мереж.
10. Імітаційне моделювання – застосування мереж Петрі.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Базова література

1. Бардачов Ю.М. Дискретна математика / Ю.М.Бардачов, Н.А.Соколова, В.Є.Ходаков – К.: Вища школа, 2002. –287 с.
2. Мурата Т. Сети Петри: Свойства, анализ, приложения / Т.Мурата / ТИИЭР, 1989. – Т. 77, № 4. – С. 41-85.
3. Нікольський Ю.В. Дискретна математика / Ю.В.Нікольський, В.В.Пасічник, Ю.М.Щербина – К.: Видавнича група BHV, 2007. – 368 с.
4. Новиков Ф.А. Дискретная математика для программистов. 2ое издание. Учебник для ВУЗов / Ф.А.Новиков – СПб.: Питер, 2006. – 368 с.
5. Питерсон Дж. Теория сетей Петри и моделирование систем: Пер. с англ. / Дж.Петерсон – М.: Мир, 1984. – 264 с
6. Таран Т. А. Основы дискретной математики. Учебное пособие / Т.А.Таран – К.: Просвіта, 1998. – с. 148.
7. Темнікова О.Л. Дискретна математика: практикум з дисципліни «Дискретна математика» для студентів спеціальності 113 «Прикладна математика» [Електронне видання] / О.Л.Темнікова – К. : КПІ ім. Ігоря Сікорського, 2018. – 88 с.

Допоміжна література

1. Гильберт Д. Основания математики. Логические исчисления и формализация арифметики / Д.Гильберт, П.Бернайс – М.: Наука, 1979. – 560 с.
2. Гиндикин С. Г. Алгебра логики в задачах / С.Г.Гиндикин – М.: Наука, 1972. – 288 с.
3. Кузнецов О. П. Дискретная математика для инженера. Учебное пособие. 6-е изд., стер. / О.П.Кузнецов – СПб.: Издательство «Лань», 2009. – 400 с
4. Нефедов В. Н. Курс дискретной математики / В.Н.Нефедов, В.А.Осипова – М.: МАИ, 1992. – 264 с.