Міністерство освіти та науки України
Київський національний університет імені Тараса Шевченка

На правах рукопису

ПУЗІКОВА АННА ВАЛЕНТИНІВНА

УДК 004.655

ТЕОРІЯ НОРМАЛІЗАЦІЇ В ТАБЛИЧНИХ БАЗАХ ДАНИХ

01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем

Дисертація на здобуття наукового ступеня
кандидата фізико-математичних наук

Науковий керівник
доктор фіз.-мат. наук,
проф. Буй Дмитро Борисович

Київ – 2016

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

БД	база даних
СУБД	система управління базами даних
НФ	нормальна форма
ФЗ	функціональна залежність
БЗЗ	багатозначна залежність
□	кінець формулювання твердження, леми або теореми, доведення
■	початок логічної частини доведення
▪	кінець логічної частини доведення
$*(R_1, R_2, ..., R_n)$	залежність з’єднання за підсхемами $R_1, R_2, ..., R_n$
Δ	множина потенційних ключів
Σ	множина обмежень
$\vdash$	відношення синтаксичного слідування
$\models$	відношення семантичного слідування
$\pi_{X,R}$	операція проекції за множиною атрибутів X таблиць схеми R
$\otimes$	операція з’єднання (природнього з’єднання)
2^X	булеан множини X
$U \mid X$	обмеження (звуження) відношення U за множиною X
$[X]_F$	замикання множини атрибутів X відносно множини ФЗ F
$[F]_\vdash$	синтаксичне замикання множини ФЗ F
$=_X$	відношення еквівалентності для рядків, які збігаються на множині атрибутів X
$[X]^{bas}_{F \cup G, R}$	базис множини атрибутів X відносно множини залежностей $F \cup G$ і схеми R
$range(s)$	множина значень рядка (функції) s

ВСТУП

Актуальність теми дослідження. В багатьох сучасних напрямках розвитку технологій баз даних (БД) в тій чи іншій мірі використовується реляційна модель даних, яка була запропонована Е. Коддом у 70-х роках XX ст.

Однією з ключових вимог, яка висувається до розробників БД, є забезпечення їх надійності, яка, в свою чергу, значною мірою залежить від правильності логічного проектування схеми БД. Проблема звільнення від відомих аномалій (оновлення, вставки, знищення) в реляційних БД, яка виникає на етапі логічного проектування внаслідок наявності різних видів обмежень (наприклад, функціональних і багатозначних залежностей), вимагає здійснення нормалізації – зведення до відповідних нормальних форм. Про потреби в автоматизації процесу нормалізації свідчать розробки CASE-засобів (Computer-Aided Software Engineering tools), наприклад, таких як ERwin[1], Vantage Team Builder (Cadre), Silverrun[2]), які здійснюють нормалізацію до третоъї нормальної форми, що, в свою чергу, вимагає залучення формальних, насамперед, математичних методів.

Теорія нормалізації спирається на теорію функціональних і багатозначних залежностей, в основі якої лежать відповідні аксіоматики та твердження про їх коректність і повноту. Аналіз наукової і методичної літератури показав, що в ній відсутні доведення згадуваних результатів, які б задовольняли стандартним вимогам до строгості математичного доведення, а також, відсутні критерії, за яких аксіоматики функціональних і багатозначних залежностей є повними.

Зв’язок роботи з науковими програмами, планами, темами. Дисертаційна робота є складовою частиною наукових робіт, які велись на кафедрі теорії та технології програмування факультету кібернетики

[1] http://erwin.com/products/detail/ca_erwin_process_modeler/
[2] http://www.silverrun.com/

Київського національного університету імені Тараса Шевченка при виконанні фундаментальної теми: "Формальні специфікації та методи розробки надійних програмних систем" (№ 0111U007052, 2011-2015 рр.).

Мета і завдання дисертаційного дослідження. *Метою* дисертаційної роботи є математичні результати щодо коректності та повноти аксіоматик функціональних і багатозначних залежностей в табличних (реляційних) БД, а також цілісний несуперечливий фрагмент математичної теорії нормалізації стосовно другої-четвертої нормальних форм.

Із огляду на мету в роботі ставляться такі *задачі*:

– побудувати математичне доведення відомого класичного результату в теорії реляційних БД про повноту аксіоматики Армстронга для функціональних залежностей, яке б задовольняло стандартним вимогам строгості та повноти математичного доведення;
– встановити критерій повноти аксіоматики ФЗ Армстронга;
– довести незалежність складових аксіоматики Армстронга в тому розумінні, що без втрати повноти з аксіоматики не можна видалити будь-яку складову;
– побудувати та дослідити алгебру функціональних залежностей;
– побудувати математичне доведення відомого класичного результату в теорії реляційних БД про повноту аксіоматики багатозначних і функціональних залежностей, яке б задовольняло традиціям встановлення повноти в аксіоматичних системах;
– встановити критерії повноти аксіоматики багатозначних залежностей та аксіоматики багатозначних і функціональних залежностей;
– побудувати фрагмент математичної теорії нормалізації щодо нормальних форм 2-4 порядків.

Об'єктом дисертаційного дослідження є табличні БД. *Предметом* дослідження є аксіоматики функціональних і багатозначних залежностей та нормалізація в табличних БД.

У роботі використовуються теоретико-множинні та логіко-алгебраїчні методи.

Наукова новизна одержаних результатів. Доведення повноти аксіоматик функціональних і багатозначних залежностей доповнено строгим математичним доведенням їх коректності; для цього були введені відношення синтаксичного і семантичного слідування для кожної з цих аксіоматик.

Вперше встановлено критерії повноти аксіоматик функціональних і багатозначних залежностей в термінах потужностей множин атрибутів та універсального домена.

Певні важливі для доведення основних результатів властивості, по-перше, замикання множини атрибутів в зазначених аксіоматиках та, по-друге, базису в аксіоматиці багатозначних і функціональних залежностей також розглянуті вперше.

Доведена незалежність складових аксіоматики Армстронга в вище зазначеному розумінні.

Побудовано алгебру функціональних залежностей − алгебраїчний аналог аксіоматики Армстронга, що дозволяє при дослідженні аксіоматики використовувати алгебраїчну мову.

Теоретичне і практичне значення одержаних результатів. Дисертація має теоретико-прикладну спрямованість. Отримані результати можуть бути застосовані в учбовому процесі та при розробці CASE-засобів, які підтримують нормалізацію.

Отримані результати були впроваджені у навчальний процес за спеціальністю "Інформатика" на факультеті кібернетики Київського національного університету імені Тараса Шевченка (нормативний курс "Композиційна семантика SQL-подібних мов").

У Кіровоградському державному педагогічному університеті імені Володимира Винниченка на кафедрі інформатики був прочитаний нормативний курс "Бази даних та СУБД", що включає результати, отримані в дисертації.

Особистий внесок здобувача. Основні результати роботи отримані здобувачем самостійно. Статті [10-13, 48-51, 54-55, 58-59, 61-62] написані у співавторстві з науковим керівником, якому належить постановка задачі дослідження, вибір методів дослідження та обговорення результатів. Із праць, виконаних зі співавторами, на захист виносяться лише результати, отримані особисто здобувачем.

Апробація результатів дослідження. Основні положення та висновки дисертаційного дослідження обговорювалися на наукових семінарах кафедри теорії та технології програмування Київського національного університету імені Тараса Шевченка.

Результати дисертаційного дослідження оприлюднені у доповідях і повідомленнях на Міжнародних та Всеукраїнських наукових конференціях, семінарах: І Міжнародному семінарі "Critical Infrastructure Safety and Security" – CrISS-Dessert'11 (Кіровоград, Україна, 2011 р.), VIII Міжнародній конференції "Теоретичні та прикладні аспекти побудови програмних систем" – TAAPSD'2011 (Ялта, Україна, 2011 р.), XVI Міжнародній конференції "Проблемы теоретической кибернетики" (Нижній Новгород, РФ, 2011 р.), X Міжнародній конференції "Теоретичні та прикладні аспекти побудови програмних систем" – TAAPSD'2013 (Ялта, Україна, 2013 р.), VII Міжнародній науково-технічній конференції "Dependable Systems, Services and Technologies" – DESSERT'2014 (Київ, Україна, 2014 р.), XVII Міжнародній конференції "Проблемы теоретической кибернетики" (Казань, РФ, 16–20 червня 2014 р.), Міжнародній науково-технічній конференції "Компьютерное моделирование в наукоемких технологиях" (Харків, Україна, 28-31 травня 2014 р.), IX Міжнародній конференції "Інтернет-Освіта-Наука-2014" – ІОН-2014 (Вінниця, Україна, 14-17 жовтня

2014 р.), XI Міжнародній науковій конференції “Теоретичні та прикладні аспекти побудови програмних систем” – TAAPSD’2014 (Київ, Україна, 15-17 грудня 2014 р.), Міжнародній науковій конференції “Сучасні проблеми математичного моделювання та обчислювальних методів” (Рівне, Україна, 19-22 лютого 2015 р.), XVII Міжнародній конференції “Дискретные модели в теории управляющих систем” (Москва и Подмосковье, РФ, 20–22 мая 2015 г.), Tenth International Conference on Dependability and Complex Systems – DepCoS-RELCOMEX (June 29 – July 3 2015, Brunów, Poland), Workshop on Foundations of Informatics – FOI-2015 (August 24-29, 2015, Chisinau, Republic of Moldova) [1210, 52, 60, 62, 13, 57, 58, 56, 50, 49, 53, 11, 10].

Публікації. Результати дисертаційного дослідження опубліковано у 19 працях. Серед них – 7 статей у наукових журналах і збірниках наукових праць [11, 48, 51, 54-55, 59, 61], з них 6 статей опубліковано у фахових виданнях, затверджених ВАК України, 1 стаття опублікована у науковому фаховому іноземному виданні [11]; 12 праць конференцій [10, 12-13, 49-50, 52-53, 56-58, 60, 62].

Структура та обсяг дисертації. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел (81 найменування) і двох додатків. Загальний обсяг дисертації становить 132 с., основний зміст викладено на 107 с. Праця містить 9 табл. та 5 рис.

РОЗДІЛ 1
ТЕОРІЯ НОРМАЛІЗАЦІЇ В РЕЛЯЦІЙНИХ БАЗАХ ДАНИХ: СУЧАСНИЙ СТАН

Теорія нормалізації в реляційних БД почала формуватись у 70-х роках минулого століття і на даний час є достатньо розвиненим розділом теорії реляційних БД. Однак, накопичених теоретичних досліджень недостатньо для задоволення потреб розробників БД; про це свідчать роботи, присвячені шляхам вирішення існуючих проблем проектування схем БД (наприклад, [19, 30, 79]) та вдосконаленню алгоритмічного апарату (наприклад, [4, 64, 65, 69]).

1.1 Логіко-історичний огляд розвитку теорії нормалізації (класичні нормальні форми 1-4 порядків)

Вперше термін «нормалізація» був застосований у 1970 р. Е. Коддом для назви процедури усунення непростих доменів [16, с. 381]. Різні підходи до інтерпретації поняття «непростого домену» впливали на інтерпретацію першої нормальної форми (1НФ), декілька означень якої наведено нижче для порівняння:

- 1975 р. – «Відношення знаходиться в 1НФ, якщо кожне значення у відношенні – тобто, кожне значення домену в кожному кортежі – є атомарним (таким, що не розкладається) елементом даних, наприклад, числом або рядком символів» [66];
- 2000 р. – «Відношення знаходиться в 1НФ, якщо домен кожного атрибута містить тільки атомарні (прості, неподільні) значення, і значення кожного атрибуту набувають тільки повних значень з цього домену» [25];

- 2005 р. – «Змінна-відношення (relvar) знаходиться у 1НФ тоді і тільки тоді, коли в кожному її допустимому значенні кожний кортеж містить тільки одне значення для кожного атрибуту» [67].

Останнє означення є більш гнучким, оскільки допускає значення складеного типу[3]. Можливість використовувати в якості елементів домену групи (groups) та відношення відстоювали автори робіт [33, 38]. Повністю формалізувати означення 1НФ не є можливим, оскільки поняття атомарності для значень домену залежить від специфіки предметної області і конкретних задач їх обробки.

У 1971 р. Е. Кодд у роботі [17] вказує на надлишковість даних та аномалії, які виникають при здійсненні операцій над відношеннями, вперше представляє концепцію функціональної залежності (ФЗ), демонструє можливість її використання для розв'язання проблем проектування БД, наводить означення другої нормальної форми (2НФ), транзитивної ФЗ та третьої нормальної форми (3НФ). У науковій літературі зустрічаються їх різні інтерпретації. Змістовне означення 2НФ використовується до цього часу: «Немає непервинного атрибута у відношенні, який би функціонально залежав від повної підмножини потенційного ключа». Надалі повна ФЗ виокремлюється в окреме поняття, яке використовується у визначенні 2НФ [67, 68, 73].

Означення транзитивної ФЗ, сформульоване Коддом у вигляді: «Нехай A, B і C різні атрибути (або різні множини атрибутів) у відношенні і нехай вони задовольняють наступним умовам: виконуються ФЗ $A \to B$, $B \to C$ та не виконується ФЗ $B \to A$. Тоді ФЗ $A \to C$ називається транзитивною залежністю», на даний час має два уточнення:

1) атрибут $C \notin A \cup B$, де A і B – множини атрибутів [73, с. 111]. Дана умова виключає з розгляду тривіальні ФЗ.

[3] Наведемо дуже симптоматичне твердження щодо атомарних значень: «Це дуже поширена помилка, що реляційна модель допускає тільки досить прості типи: числа, рядки, можливо, дати та час, і не більше. Скоріше, відношення можуть мати атрибути будь-якого роду, [...] ці типи повинні бути, і насправді, можуть бути, такими складними, як нам подобається [...]» [20].

2) $C \not\subseteq B$, $B \not\subseteq A$, де A, B і C – множини атрибутів [68, с. 24]. Цим також виключаються тривіальні ФЗ в посилках $A \to B$ і $B \to C$.

В залежності від використовуваних уточнень означення 3НФ сформульоване Коддом у вигляді «Кожний непервинний атрибут не є транзитивно залежним від будь-якого потенційного ключа у відношенні», на даний час має два варіанти.

1) Відношення знаходиться у 3НФ, якщо воно знаходиться у 1НФ і не містить транзитивних залежностей (з вище наведеним уточненням 1)) непервинних атрибутів від потенційних ключів [73]. Додатково доводиться теорема, яка стверджує, що з знаходження відношення у 3НФ випливає знаходження цього відношення у 2НФ [73, с. 111].

2) Відношення знаходиться у 3НФ, якщо воно знаходиться у 2НФ і не містить транзитивних залежностей (з вище наведеним уточненням 2)) непервинних атрибутів від потенційних ключів (див., наприклад, [67]).

Строге математичне викладення концепції ФЗ та теорії нормальних форм (НФ) станом на початок 80-х років виконано у монографії В. П. Дрібаса [68], зокрема, детально розглянуті еквівалентні умови, за яких множина атрибутів відношення є потенційним ключем, досліджені можливі взаємозв’язки між транзитивними та неповними ФЗ, сформульовані означення посилених та оптимальних 2НФ і 3НФ, поняття НФ поширене на набір проекцій відношень, які задовольняють певним умовам.

Відкриття аксіом та правил виведення ФЗ із заданої множини ФЗ [24] та побудова аксіоматики Армстронга для ФЗ [2] дали можливість розробити алгоритми обчислення так званого канонічного покриття для заданої множини ФЗ та замикання для множини атрибутів [7, 65, 73, 78].

Одним із способів приведення відношення до 3НФ є декомпозиція, обґрунтуванням якої стала теорема Хеза (Heath) [31] про виконуваність декомпозиції без втрат відношення $R(A, B, C)$ на проекції $R_1 = \pi_{(A,B),R}$ та $R_2 = \pi_{(A,C),R}$ за умови виконання ФЗ $A \to B$. До проблем декомпозиції належить залежність проекцій (за термінологією Ріссанена (Rissanen) [42]), яка може стати причиною аномалій. Дана проблема розв'язана у статті А. Філіповича [79], який розглянув взаємні ФЗ (ВФЗ) і умовні взаємні ФЗ (УВФЗ) та дослідив їх властивості, запропонував алгоритм виявлення ВФЗ, поняття взаємно-незалежної НФ (ВННФ), яку можна розглядати як синонім ациклічної БД, та спосіб зведення до неї.

Інший спосіб запропонував Бернштейн (Bernstein), який використав альтернативне означення 3НФ (відношення знаходиться у 3НФ, якщо кожний атрибут, який транзитивно залежить від ключа, є первинним атрибутом) для побудови алгоритму синтезу повної схеми БД у 3НФ для заданої множини ФЗ [8].

Специфіка різних підходів до задачі проектування схеми реляційної БД викликана відмінностями у формальних визначеннях еквівалентності та критеріїв якості схеми [6].

Відомими класичними алгоритмами зведення схеми відношення до 3НФ є алгоритми Ульмана [78], Делобеля-Гейсі (Delobel-Gasey) [24], результатами яких не завжди є схема у 3НФ, Берштейна [8], Іслура (Isloor) [32], Неклюдової-Цаленка [74], який дає кількісно оптимальну схему БД; оригінальний алгоритм зведення схеми відношення до 3НФ через побудову кільцевих покриттів запропонував Мейер [73]. Переваги та недоліки більшості з вказаних алгоритмів, а також їх відповідність різним визначенням еквівалентності реляційних схем розглянуті у монографії Дрібаса [68]. Пошук ефективних алгоритмів розв'язання задачі синтезу оптимальної схеми БД у 3НФ продовжується і сьогодні (наприклад, [64, 65, 69]).

Недоліки 3НФ були розглянуті і враховані у роботі Хеза [31] при формулюванні означення посиленої 3НФ та, пізніше, у роботі Кодда [18] (інша назва означення – нормальна форма Бойса-Кодда (НФБК)). Одним з перших відомих алгоритмів зведення «майже» до НФБК є алгоритм Берштейна [8], який дозволяє усувати транзитивні залежності первинних атрибутів від ключів, що не містять ці атрибути. Більш пізні алгоритми наводяться, наприклад, у роботах [35, 4].

З введенням у розгляд багатозначних залежностей (БЗЗ) [45] Р. Фагін (Fagin) досліджує їх властивості та визначає нову четверту НФ (4НФ) [27, с. 267]. Зокрема, у зазначеній роботі доведена важлива для теорії нормалізації теорема: «БЗЗ $X \rightarrow\rightarrow Y$ виконується для відношення $R(X,Y,Z)$ тоді і тільки тоді, коли відношення R дорівнює з'єднанню його проекцій $R_1(X,Y)$ і $R_2(X, Z)$», та показано, що якщо схема відношення знаходиться у 4НФ, то вона знаходиться у НФБК. Строгий та повний набір правил виведення для БЗЗ, правила, які пов'язують ФЗ та БЗЗ, їх коректність та повнота представлені у статті [5], незалежність побудованої системи аксіом обговорюється у роботі Мендельзона (Mendelzon) [39], а проблеми її повноти та ненадлишковості – у статті Біскапа (Biskap) [9].

Детальне викладення результатів з теорії залежностей та теорії нормалізації в реляційних БД здійснено у монографії Мейера [73]. Розділи 4-5 роботи присвячені розгляду ФЗ та покриття множини ФЗ, правилам виведення ФЗ з заданої множини ФЗ, проблемі збігання істинності та вивідності (тобто, семантичного та синтаксичного слідувань), побудові алгоритмів перевірки того, що деяка логічна умова є наслідком заданої множини умов, побудові ациклічних графів виведення. У наступних розділах розглянуті: проблеми аномалій та надлишковості даних, 2-3 НФ та НФБК, проблеми нормалізації через декомпозицію та алгоритм синтезу схеми реляційної БД (розділ 6); БЗЗ та залежності з'єднання (ЗЗ), аксіоматика для БЗЗ, 4-5 НФ (розділ 7); метод перевірки синтаксичного слідування ФЗ або ЗЗ із заданої множини ФЗ та ЗЗ (розділ 8); задача проектування реляційної схеми (розділ 9).

ЗЗ та аномалії, які вони викликають, були розглянуті у статті Ріссанена (Rissanen) [42] та досліджені у роботах [1, 23]. Зокрема, у статті [1] наводиться алгоритм перевірки, чи буде результат з'єднання проекцій відношення задовольняти умові з'єднання без втрат, який працює за умови, що усі залежності є функціональними. У роботі [23] доводиться важлива для теорії нормалізації теорема: «Якщо $R = \langle U, G \rangle$, де G – система твірних структури БЗЗ, то для U_1 та U_2, таких, що $U_1 \bigcup U_2 = U$, схеми $\{R_1(U_1), R_2(U_2)\}$ володіють властивістю з'єднання без втрат тоді і тільки тоді, коли БЗЗ $U_1 \bigcap U_2 \to\to U_1$ (або $U_1 \bigcap U_2 \to\to U_2$) належить до G^+». Крім того наводиться алгоритм перевірки, чи задовольняє деяка декомпозиція умові декомпозиції без втрат.

З результату про відсутність повної скінченної множини правил виведення для ЗЗ, який наводиться у роботі С. Петрова [75], випливає неможливість побудови скінченної повної та коректної аксіоматики для ЗЗ.

Ґрунтовне математичне викладення теорії залежностей та окремі питання теорії нормалізації схеми реляційної БД здійснено у главах 5-6 монографії М. Цаленка [80]. Зокрема, розглянуто ідею про відповідності Галуа між теоріями та їх моделями, специфіка якої для БД полягає в тому, що, як правило, для довільної множини формул, замкненої відносно логічного слідування, існує одна модель, яка повністю визначає дану множину – БД Армстронга. Для частини семантики предметної області, яка описується ФЗ та БЗЗ, побудовано скінченну дедуктивну систему з алгоритмічно розв'язувальною проблемою синтаксичного слідування; встановлюється відсутність існування скінченної дедуктивної системи для вбудованих БЗЗ (ВБЗЗ). Описані два методи розв'язання алгоритмічних задач дедуктивного виведення: перший ґрунтується на застосуванні апарату булевих функцій, другий – на використанні методу табло.

1.2 Аналіз еквівалентності означень класичної п'ятої нормальної форми

Концепцію проективно-з'єднувальної НФ (Projection-Join Normal Form – PJ/NF) або класичної п'ятої НФ (5НФ) Фагін розглянув у статті [28], запропонувавши два означення. В першому – використовується алгоритм приналежності (Membership algorirhm), який працює наступним чином: на вхід подаються множина потенційних ключів $\Delta = \{K_1, K_2, ..., K_m\}$ відношення із множиною атрибутів R реляційної схеми[4] та ЗЗ $*(R_1, R_2, ..., R_n)$. Ініціюється множина підсхем $S = \{R_1, R_2, ..., R_n\}$. Поки для деякого K_i, $1 \le i \le m$, виконується включення $K_i \subseteq R_j \cap R_k$, де $1 \le j, k \le n$, в множині S входження множини R_j замінюється на множину $R_j \cup R_k$, а входження R_k знищується. Якщо на виході множина S «співпадає» з схемою R, тобто $S = \{R\}$, то алгоритм приналежності видає значення «*True*». Приймається без доведення (на прикладі), що алгоритм видає результат «*True*» тоді і тільки тоді, коли ЗЗ $*(R_1, R_2, ..., R_n)$ логічно (семантично) слідує з множини ключів $\Delta = \{K_1, K_2, ..., K_m\}$. Таким чином, згідно [28], наступні означення є еквівалентними:

1) «схема відношення знаходиться у PJ/NF, якщо вона знаходиться у 1НФ і для множини ключів Δ та кожної ЗЗ $\sigma \in \Sigma$ алгоритм видає результат «*True*»;

2) «схема відношення знаходиться у PJ/NF, якщо вона знаходиться у 1НФ і кожна ЗЗ $\sigma \in \Sigma$ є логічним (семантичним) наслідком множини ключів Δ $(\Delta \vDash \sigma)$».

---

[4] Зауважимо, що в роботі [28] під реляційною схемою розуміється пара $\langle R, \Sigma \rangle$, де R – множина атрибутів, Σ – множина обмежень, яка складається з ФЗ, БЗЗ та ЗЗ і є замкненою відносно операції логічного (семантичного) слідування. Також нагадаймо, що в ЗЗ вигляду $*(R_1, R_2, ..., R_n)$ множини атрибутів R_i (підсхеми) називаються компонентами ЗЗ.

Перше означення по суті є синтаксичним, оскільки застосовується до схеми відношення, а друге – семантичним, оскільки в ньому неявно використовується поняття моделі множини ключів Δ, причому на цій моделі повинна виконуватись кожна залежність з'єднання $\sigma \in \Sigma$. Проаналізуємо детальніше, чи дійсно ці означення є еквівалентними.

Лема 1.1. Якщо таблиця t знаходиться у PJ/NF за означенням 1), то вона знаходиться у PJ/NF за означенням 2). □

Доведення. Нехай таблиця t є моделлю реляційної схеми $\langle R, \Sigma \rangle$, яка знаходиться у PJ/NF за означенням 1). Тоді для довільної ЗЗ $\sigma \in \Sigma$, представленої у вигляді $*(R_1, R_2, \ldots, R_n)$, на деякому кроці виконання алгоритму приналежності для деякого потенційного ключа $K_i \in \Delta$, $1 \le i \le m$, знайдуться такі компоненти R_j і R_k, $1 \le j, k \le n$, що виконується включення $K_i \subseteq R_j \cap R_k$. Звідси за пунктом 2 твердження 2.10.4 монографії [76, с. 74] випливає, що виконується рівність $\pi_{R_j \cup R_k}(t) = \pi_{R_j}(t) \otimes \pi_{R_k}(t)$.

Вище π_X позначає операцію проекції за множиною X таблиці схеми R, $\otimes$ – операцію природного з'єднання (natural join, за термінологією [76, с. 32] – з'єднання) таблиць.

Продовжуючи застосовувати вказаний пункт для кожного наступного включення вигляду $K_i \subseteq R_j \cap R_k$, $1 \le i \le m$, в результаті виконання алгоритму приналежності з урахуванням властивостей комутативності та асоціативності з'єднання [76] отримаємо рівність:

$$\pi_{R_1 \cup \ldots \cup R_n}(t) = \pi_{R_1}(t) \otimes \pi_{R_2}(t) \otimes \ldots \otimes \pi_{R_n}(t),$$

оскільки $R = R_1 \cup \ldots \cup R_n$, то $\pi_R(t) = t = \pi_{R_1}(t) \otimes \pi_{R_2}(t) \otimes \ldots \otimes \pi_{R_n}(t)$.

Вище використали властивість проекції $\pi_X(t) = t$ за умови $R \subseteq X$, де R – схема таблиці t (пункт 2 твердження 2.4.1 з [76]).

Звідси випливає, що ЗЗ $\sigma \in \Sigma$ виконується на таблиці t, а отже реляційна схема $\langle R, \Sigma \rangle$ знаходиться у PJ/NF за означенням 2). □

Покажемо, що обернена імплікація не виконується; тобто, якщо таблиця t знаходиться у PJ/NF за означенням 2), то вона необов'язково знаходиться у PJ/NF за означенням 1). Для цього достатньо розглянути наступний контрприклад.

Приклад 1.1. Нехай над множиною атрибутів $R=\{A,B,C\}$ з потенційним ключем $\{A\}$ задана ЗЗ $*(ABC,C)$. Очевидно, що на довільній таблиці $t(R)$, яка є моделлю множини ключів $\Delta=\{\{A\}\}$ вказана ЗЗ виконується, оскільки вона є тривіальною[5]. Але алгоритм приналежності видасть результат «*False*».

Таким чином, доведено наступне твердження.

Твердження 1.1. Якщо схема відношення (таблиці) знаходиться у PJ/NF за означенням 1), то вона знаходиться у PJ/NF за означенням 2). Обернена імплікація в загальному випадку не виконується. □

Отже, можемо зробити висновок, що означення 1) для PJ/NF є більш сильним, ніж означення 2).

Автори роботи [21] Дейт і Фагін розглянули еквівалентність 3НФ-PJ/NF за означенням 2) в особливих випадках, коли потенційні ключі (ключ) є простими. Зокрема, показано що, якщо схема відношення R знаходиться в 3НФ і кожний ключ складається з одного атрибута, то R знаходиться у PJ/NF за означенням 2); якщо схема відношення R знаходиться у НФБК і деякий ключ складається з одного атрибута, то R знаходиться у 4НФ.

1.3 Некласичні нормальні форми

Означення класичних НФ і супутніх понять (наприклад, потенційного ключа) в реляційних БД застосовуються до одного відношення [16, 67],

---

[5] Нагадаємо, що ЗЗ є тривіальною, якщо один з її компонентів R_i співпадає з схемою R. Тривіальні ЗЗ виконуються на довільній таблиці. Це випливає з рівностей $\pi_R(t)\otimes\pi_X(t)=t\otimes\pi_X(t)=t$, де R – схема t; доведення випливає з властивостей з'єднання (підрозділ 2.7, с. 49-56 з [7676]).

схеми одного відношення [27-28, 73, 78], таблиці [76]. Алгоритм зведення відношення до певної НФ повинен задовольняти вимогам (принципам) задачі проектування схеми БД. Одним з важливих принципів розробки схеми БД є принцип представлення [6], який включає в себе можливість відновлюваності початкового відношення r_0 зі схемою R_0 (множиною атрибутів) з відношень $r_1, r_2, ..., r_n$ зі схемами $R_1, R_2, ..., R_n$ відповідно, де r_i – проекції відношення r_0 на схеми R_i для $i = \overline{1, n}$; іншими словами, вимагається виконання декомпозиції без втрат $r_0 = \otimes_{i=1}^{n} r_i$.

Передумовою виникнення пропозицій щодо покращення існуючих класичних НФ 2-3 порядків та НФ Бойса-Кодда (НФБК) є те, що ні декомпозиція (див., наприклад, [79]), ні деякі алгоритми синтезу (наприклад, алгоритм Берштейна [8]) не забезпечували виконання принципу представлення для відношення з довільно заданими схемою і множиною ФЗ.

Так, якщо на вхід алгоритма синтезу Берштейна подається реляційна схема $R_0 = \{A, B, C, D\}$ із множиною ФЗ $\{A \to B, C \to D\}$ (приклад з монографії Дрібаса [68, с. 44]), результатом виконання алгоритму є схеми $R_1 = \{A, B\}$ і $R_2 = \{C, D\}$, кожна з яких знаходиться у 3НФ, але порушується вимога виконання декомпозиції без втрат, оскільки з’єднанням відповідних відношень r_1 і r_2 є їх декартове з’єднання.

Намагаючись уникнути подібної проблеми, Дрібас поширює означення НФ 2-3 порядків та НФБК на набір проекцій відношень, наприклад: «3НФ відношення $r(R)$ – це або дане відношення, якщо воно знаходиться у 2НФ і не містить транзитивних залежностей непервинних атрибутів від потенційних ключів, або набір яких-небудь його проекцій, кожна з яких задовольняє вказаним умовам, причому дане відношення повинно відновлюватись через природнє з’єднання (natural join) цих проекцій» [68, с. 25].

Іншу спробу поєднати вимогу виконання декомпозиції без втрат з означенням 3НФ запропонували Лінг (Ling), Томпа (Tompa) і Камеда (Kameda), які розробили концепцію покращеної 3НФ (An Improved Third Normal Form) та алгоритм зведення до неї [36].

Відомо, що алгоритм синтезу Берштейна моделює схеми відношень, які знаходяться «майже» в НФБК. Маючи на меті охарактеризувати форму цих відношень, Заніоло запропонував НФ, яка визначається елементарним ключем (Elementary Key Normal Form – EKNF) [46]. Автор сформулював більш строге означення 3НФ: «Відношення знаходиться у 3НФ за означенням, якщо для кожної нетривіальної ФЗ $X \to A$:

a) множина атрибутів X є суперключем відношення із схемою R, або

b) атрибут A належить деякому потенційному ключу»,

та звернув увагу на різницю між означеннями 3НФ і НФБК – остання не містить умову b).

В означенні EKNF використовуються додаткові поняття елементарної ФЗ та елементарного ключа:

– ФЗ $X \to A \in F$ (де F – множина ФЗ) називається елементарною, якщо $A \notin X$ і для кожної множини атрибутів $X' \subset X$ ФЗ $X' \to A$ не належить $[F]$ (де $[F]$ – замикання множини ФЗ); зауважимо, що за термінологією [68] така ФЗ $X \to A$ є повною;

– ключ X у відношенні схеми R є елементарним, якщо для деякого атрибута $A \in R$ ФЗ $X \to A$ є елементарною; наприклад, для схеми відношення $R = \{A, B, C\}$ із заданою множиною ФЗ $\{AB \to C, C \to A\}$ ключ $\{A, B\}$ є елементарним, а ключ $\{B, C\}$ – ні.

В означенні EKNF зберігається умова (a) та посилюється умова (b), вказані вище: «Відношення із схемою R знаходиться у EKNF за означенням, якщо для кожної нетривіальної ФЗ $X \to A$:

a) X є суперключем відношення, або

b) атрибут A належить деякому елементарному ключу».

На зв'язок EKNF з класичними НФ вказують імплікації НФБК$\Rightarrow$EKNF$\Rightarrow$3НФ, які випливають безпосередньо з наведених вище означень.

Як вже згадувалось, схему відношення у EKNF синтезує алгоритм Берштейна; доведення цього факту спирається на теорему про те, що кожен ключ, який отримується в результаті виконання вказаного алгоритму, є елементарним [46].

Перед тим, як перейти до наступного блоку некласичних НФ, пов'язаних із означеннями PJ/NF або класичної 5НФ, розглянемо детальніше визначення 5НФ, сформульоване Дейтом [67], та попередню версію 5НФ у монографії Мейера [73, с. 151]. Наведемо обидва визначення:

1) змінна-відношення знаходиться в 5НФ (PJ/NF) за означенням, якщо кожен компонент кожної нетривіальної ЗЗ є суперключем (Дейт);

2) схема відношення R знаходиться у попередній PJ/NF за означенням, якщо кожна ЗЗ $*(R_1, R_2, ..., R_n)$ $(R_1 \cup R_2 \cup ... \cup R_n = R)$, яка виводиться (синтаксично слідує) з множини обмежень Σ, є або тривіальною, або її компоненти є суперключами (Мейер).

Використання поняття суперключа в цих означеннях обґрунтовується наступною лемою.

Лема 1.2. Якщо схема відношення $\langle R, \Sigma \rangle$ знаходиться у PJ/NF за означенням 1), то кожен компонент довільної ЗЗ $\sigma \in \Sigma$ є суперключем. □

Доведення проводиться безпосередньо. □

Наслідок 1.1. Якщо схема відношення $\langle R, \Sigma \rangle$ знаходиться у PJ/NF за означенням 1), то кожен атрибут (точніше кажучи, відповідний сінглітон) з множини R є суперключем[6].

Доведення. Нехай схема відношення $\langle R, \Sigma \rangle$ знаходиться у PJ/NF за означенням 1) і нехай ЗЗ $\sigma \in \Sigma$ представлена у вигляді $*(R_1, R_2, ..., R_n)$. Далі

[6] Нагадаємо означення суперключа: «Суперключем є множина атрибутів $X \subseteq R$, яка містить ключ. Кожен ключ є також суперключем» [8, с. 279].

враховуємо, що множина ФЗ і ЗЗ Σ є замкненою відносно операції логічного (семантичного) слідування і виконується включення $\{\sigma \mid \Sigma \vDash \sigma\} \supseteq \{\varphi \mid \Sigma \vdash \varphi\}$, яке випливає з результату про коректність відповідної аксіоматизації для ЗЗ [75]. Вище $\vdash$ – відношення синтаксичного слідування. Звідси випливає, що $\forall \varphi (\Sigma \vdash \varphi \Rightarrow \varphi \in \Sigma)$.

Нехай $A \in R$, тоді ЗЗ $*(R_1, R_2, ..., R_n, \{A\})$, яка отримана за правилом виведення $\dfrac{*(R_1, R_2, ..., R_n)}{*(R_1, R_2, ..., R_n, X)}$, де $\varnothing \neq X \subseteq R$, належить множині Σ, а отже, за лемою 1.2 одноелементна множина $\{A\}$ є суперключем. □

Наведемо приклад з [73], на якому Мейер демонструє нееквівалентність означень попередньої PJ/NF і класичної PJ/NF за означенням 1):

«Нехай задані схема відношення $R = \{A, B, C\}$ і множина обмежень, яка містить ФЗ і ЗЗ: $\Sigma = \{A \to BC, C \to AB, *(AB, BC)\}$. Оскільки множини $\{A, B\}$ і $\{B, C\}$ є суперключами, то схема відношення знаходиться у підготовчій PJ/NF, але вона не знаходиться у класичній PJ/NF, оскільки $\{A, B\} \cap \{B, C\} = \{B\}$ не є ключем».

Насправді, множина $\{B\}$ в даному прикладі є ключем; це випливає з наступної леми, ідея якої неявно викладена в роботі Вінсента (Vincent) [43].

Лема 1.3. Якщо на таблиці t схеми R для множин атрибутів R_1, R_2 таких, що $R_1 \cup R_2 = R$ виконується нетривіальна ЗЗ $*(R_1, R_2)$ і R_1, R_2 є суперключами, то $R_1 \cap R_2$ є суперключем. □

Доведення. Нехай на таблиці t виконується нетривіальна ЗЗ $*(R_1, R_2)$, де R_1, R_2 – суперключі, і нехай $R_1 \cap R_2 = X$. Розглянемо випадки:

– $X = \varnothing$; тоді з'єднанням проекцій $\pi_{R_1}(t)$ і $\pi_{R_2}(t)$ є їх декартове з'єднання, моделями такої ЗЗ можуть бути порожня таблиця ($t_\varnothing$) або однорядкова таблиця, для яких множина $X = \varnothing$ є єдиним ключем [62]; ▪

– $X \neq \varnothing$; тоді за теоремою 1 Фейгіна [27, с. 266] з виконання ЗЗ $*(R_1, R_2)$ випливає виконання БЗЗ $X \to\to R_1 \setminus X$. Оскільки R_2 є суперключем, то виконується ФЗ $R_2 \to R_1 \setminus X$. Звідси за спільним правилом для ФЗ і БЗЗ $\dfrac{X \to\to Z, Y \to Z'}{X \to Z'}$, де $Z' \subseteq Z$, $Y \cap Z = \varnothing$, випливає виконання ФЗ $X \to R_1 \setminus X$. За правилом поповнення для ФЗ маємо виконання ФЗ $X \to R_1$; звідси за правилом транзитивності з урахуванням ФЗ $R_1 \to R$ випливає, що ФЗ $X \to R$ також виконується. ▪

Отже, $R_1 \cap R_2$ є суперключем. □

Лема 1.4. Якщо відношення знаходиться у PJ/NF за означенням 1) Фейгіна, то воно знаходиться у попередній PJ/NF за означенням Мейера та 5НФ за означенням Дейта. □

Доведення випливає безпосередньо з леми 1.2. □

Вінсент у роботі [43] доводить лему про те, що відношення знаходиться у 5НФ (за означенням Дейта) або попередній PJ/NF (за означенням Мейера) тоді і тільки тоді, коли кожен атрибут множини R є ключем.

Вкажемо суттєвий недолік означень Дейта і Мейера, а також означення 1) для PJ/NF. В формулюванні означення 2) для PJ/NF Фейгін дотримується принципу «включення» НФ вищих порядків в НФ нижчих порядків (див. рис. 1.1).

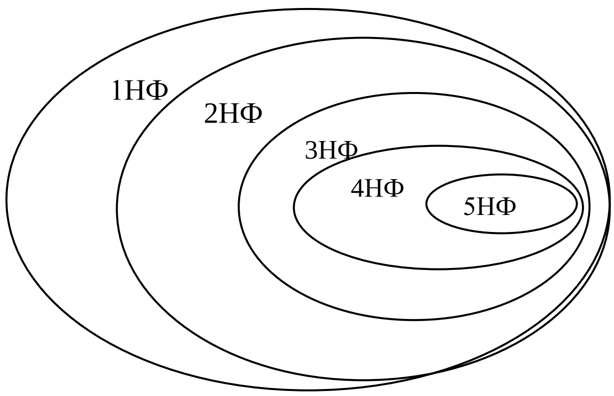

Рис. 1.1 Співвідношення між НФ.

Нагадаємо, що в роботі [28] Фейгін сформулював означення класичних НФБК та 4НФ в семантичному вигляді та довів їх еквівалентність з відповідними класичними означеннями НФ, сформульованими в синтаксичній формі. Так, за означенням Фейгіна, «схема відношення знаходиться у НФБК (4НФ), якщо вона знаходиться у 1НФ і кожна ФЗ (БЗЗ) $\sigma \in \Sigma$ є логічним наслідком множини ключів Δ ($\Delta \models \sigma$)». З того, що ФЗ є окремим випадком БЗЗ, яка в свою чергу є окремим випадком ЗЗ, випливають імплікації: PJ/NF (за означенням 2))$\Rightarrow$4НФ$\Rightarrow$НФБК. Звідси очевидно, що означення 2) PJ/NF еквівалентно означенню 4НФ за умови, що усі ЗЗ є БЗЗ.

Синтаксичні форми означень Дейта і Мейера як і означення 1) для PJ/NF не є узагальненням 4НФ, тобто, не еквівалентні означенню 4НФ, якщо кожна ЗЗ у множині залежностей є БЗЗ [43]. Скористаємось прикладом, наведеним у [43].

Приклад 2. Нехай $R = \{A, B, C\}$ і нехай $\Sigma = \{A \to B, A \to C\}$. Очевидно, що реляційна схема $\langle R, \Sigma \rangle$ знаходиться у 4НФ, але не знаходиться у 5НФ (за означеннями Дейта або Мейера), оскільки за правилом виведення $\dfrac{*(R_1, R_2, ..., R_n)}{*(R_1, R_2, ..., R_n, X)}$, де $\varnothing \neq X \subseteq R$, з ЗЗ $*(AC, AB)$ синтаксично слідує ЗЗ $*(AC, AB, C)$, компонент C якої не є ключем.

Враховуючи недоліки означень попередників Вінсент пропонує зменшити множину ЗЗ, які розглядаються, і запроваджує спрощену 5НФ (reduced-5NF – 5NFR) [43]. «Нехай Σ – множина ФЗ і ЗЗ, причому кожна ЗЗ $*(R_1, R_2, ..., R_p)$ задовольняє умовам:

1. $R_1 \cup R_2 \cup ... \cup R_p = R$, де R – схема відношення;

2. $\forall R_i, 1 \leq i \leq p$, $*(R_1, R_2, ..., R_{i-1}, R_{i+1}, ..., R_p)$ або не належить Σ^+, або $R_1 \cup R_2 \cup ... \cup R_{i-1} \cup R_{i+1} \cup ... \cup R_p \neq R$, де Σ^+ – множина усіх залежностей, які слідують (синтаксично) з множини Σ.

Пара $\langle R, \Sigma \rangle$ знаходиться у 5NFR за означенням, якщо ліва частина кожної її нетривіальної ФЗ є суперключем і для кожної ЗЗ $\sigma \in \Sigma$ кожен компонент σ є суперключем».

Альтернативне означення запропонував Норман, який розглянув поняття нескоротної ЗЗ [41]. За Норманом ЗЗ $*(R_1, R_2,...,R_n)$ є нескоротною, якщо не існує підмножини $\{R_{i_1}, R_{i_2},...,R_{i_s}\} \subset \{R_1, R_2,...,R_n\}$, такої, що ЗЗ $*(R_{i_1}, R_{i_2},...,R_{i_s})$ належить Σ^+. Тоді, реляційна схема $\langle R, \Sigma \rangle$ знаходиться у SKNF за означенням, якщо кожен компонент кожної нескоротної ЗЗ в R є суперключем [41].

Лема 1.5. Якщо схема відношення (таблиці) знаходиться у попередній PJ/NF за означенням Мейера, то воно знаходиться у 5NFR за означенням Вінсента та у SKNF за означенням Нормана. □

Доведення випливає безпосередньо з означень. □

У роботі [44] Вінсент запропонував одразу дві НФ: НФ, яка повністю визначається ключем (ключами) (Key-Complete Normal Form – KCNF) і надлишково-вільну НФ (Redandancy-free normal form – RFNF).

В означенні KCNF використовується поняття залежності з’єднання, яка повністю визначається ключами (Key-Complete Join Dependences): «Нехай множина Σ складається з ФЗ і ЗЗ. ЗЗ $*(R_1, R_2,...,R_p)$ повністю визначається ключами, якщо об’єднання тих її компонентів R_i, $1 \le i \le p$, які є суперключами, співпадає з R».

Розглянемо приклад з [44]. Якщо $R = \{A, B, C, D\}$ і $\Sigma = \{A \to BCD,$ $B \to ACD,$ $*(AB, BC, CD),$ $*(AB, BC, ACD)\}$, то множини $\{A\}$ і $\{B\}$ є потенційними ключами. ЗЗ $*(AB, BC, ACD)$ повністю визначається ключами, оскільки всі її компоненти є суперключами, а їх об’єднання $\{AB\} \bigcup \{BC\} \bigcup \{ACD\} = R$; тоді як ЗЗ $*(AB, BC, CD)$ не повністю визначається ключами, оскільки суперключами є компоненти $\{AB\}$ і $\{BC\}$, об’єднання яких $\{AB\} \bigcup \{BC\} \neq R$.

За означенням [44] пара $\langle R, \Sigma \rangle$ знаходиться у KCNF, якщо «ліва частина кожної нетривіальної ФЗ $\varphi \in \Sigma$ є суперключем, а кожна ЗЗ $\sigma \in \Sigma$ повністю визначається ключами».

Положення KCNF і 5NFR по відношенню до інших НФ встановлюють імплікації PJ/NF (за означенням 1))$\Rightarrow$5NFR$\Rightarrow$KCNF$\Rightarrow$4НФ [44].

В означенні RFNF використовується поняття надлишковості для входження значення змінної: «Входження значення змінної $t(A)$ кортежа t відношення r є надлишковим, якщо в результаті заміни $t(A)$ на довільне значення a', таке, що $t(A) \neq a'$, отримане відношення r' не належить множині $SAT(\Sigma)$, де $SAT(\Sigma)$ – множина екземплярів (моделей) множини Σ, яка складається з ФЗ і ЗЗ». Відповідно визначається надлишкове відношення: «Відношення r є надлишковим, якщо існує входження $t(A)$ значення змінної в r, яке є надлишковим». За означенням [44]: «Пара $\langle R, \Sigma \rangle$ знаходиться у надлишково-вільній нормальній формі, якщо не існує відношення $r(R) \in SAT(\Sigma)$, яке є надлишковим».

Доводиться теорема про еквівалентність RFNF та 4НФ у випадку, коли множина обмежень Σ складається лише з ФЗ і БЗЗ. Для випадку, коли об'єднання компонентів $R_1,...,R_p$ кожної ЗЗ $\sigma \in \Sigma$ співпадає з R, доводиться теорема про еквівалентність RFNF і KCNF.

Описаний підхід до нормалізації відношення в термінах ненадлишковості кортежів привернув увагу Дарвіна (Darwen), Дейта і Фейгіна. У 2012 р. вони розробили НФ з необхідними кортежами (Essential Tuple Normal Form – ETNF), яка визначається в термінах ФЗ і ЗЗ та займає проміжне положення між 4НФ та PJ/NF (за означенням 1)) [19]. В означенні ETNF використовуються додаткові означення частково надлишкового кортежу, логічного слідування кортежу з множини кортежів та повністю надлишкового кортежу:

– кортеж t екземпляра r схеми R називається частково надлишковим, якщо існує кортеж $t' \in r$, $t' \neq t$, що для нетривіальної ФЗ $X \rightarrow A \in \Sigma$ виконується рівність $t[X] = t'[X]$ (тобто, в кортежах t та t' співпадають значення всіх атрибутів з множини X);

– кортеж t логічно слідує з множини кортежів S (відносно схеми $\langle R, \Sigma \rangle$), якщо кожний екземпляр r схеми $\langle R, \Sigma \rangle$, який містить усі кортежі з множини кортежів S, також містить кортеж t;
– кортеж t екземпляра r схеми $\langle R, \Sigma \rangle$ називається повністю надлишковим, якщо існує множина кортежів S в r причому $t \notin S$, така, що з S логічно слідує t (відносно $\langle R, \Sigma \rangle$).

В термінах часткової надлишковості означення НФБК формулюється так: «Схема відношення R знаходиться у НФБК за означенням, якщо не існує її екземпляра, який би містив частково надлишковий кортеж».

Поняття часткової та повної надлишковості кортежу є спеціальними випадками поняття надлишковості [44].

Згідно означення реляційна схема $\langle R, \Sigma \rangle$ знаходиться в кортеже-необхідній НФ (Essential Tuple Normal Form), якщо кожний кортеж кожного екземпляра схеми $\langle R, \Sigma \rangle$ не є частково чи повністю надлишковим.

На положення ETNF по відношенню до інших НФ вказують імплікації: PJ/NF (за означенням 1))$\Rightarrow$SKNF$\Rightarrow$RFNF$\Rightarrow$ETNF$\Rightarrow$4НФ, причому зворотні імплікації не виконуються. Обговорюються прості умови, за яких схема відношення $\langle R, \Sigma \rangle$ буде знаходитись у ETNF, а саме: 1) схема відношення $\langle R, \Sigma \rangle$ повинна бути зведена до НФБК; 2) деякий ключ повинен складатись з одного атрибута.

Окремим видом ЗЗ є вкладені залежності (ВЗ), для яких пропонувались інші НФ. Ці НФ в даній роботі не розглядаються. Зауважимо лише, що дослідженню вкладених залежностей, їх властивостей та взаємодії з ФЗ присвячено багато робіт. Зокрема, аксіоматика для ВЗ викладена в [14]; відомими є НФ Ніколаса (Nicolas) для вкладених ЗЗ [40], Inclusion Normal Form та алгоритми зведення до неї Лінга (Ling) і Гоха (Goh) [37], Inclusion Dependency Normal Form Левена (Levene) [34].

Наостанок розглянемо доменно-ключову НФ (Domain-Key Normal Form – DK/NF або ДКНФ), запропоновану Фейгіним у роботі [26]. Концепція DK/NF базується на поняттях залежності ключа та залежності домена: «Відношення знаходиться у DK/NF за означенням, якщо кожне обмеження з його схеми є логічним наслідком з об'єднання множин залежності ключа та залежності домена». Доводиться, що реляційна схема знаходиться у DK/NF тоді і тільки тоді, коли немає аномалій вставки та знищення (означення яких формулюються у роботі); проводиться аналогія між класичними та альтернативними означеннями НФБК, 4НФ та PJ/NF (за означенням 2), розглянутими у роботі [28], пропонуються їх модифікації з урахуванням комбінаторного ефекту розмірів обмежуючого домену та показується, що DK/NF включає в себе ці модифікації; частковим є випадок, коли потужність усіх доменів є нескінченною, тоді DK/NF включає в себе НФБК, 4НФ та PJ/NF (за означенням 2) у класичному виді. Але під обмеженнями в означенні ДКНФ розуміються не тільки ФЗ, БЗЗ та ЗЗ, означення яких використовуються у попередніх НФ, а й усі, які можна записати у вигляді висловлювання логіки предикатів 1-го порядку. У роботі [29] пропонується форма запису для усіх відомих на той час залежностей:

$$(\forall x_1..x_m)((A_1 \wedge ... \wedge A_n) \Rightarrow \exists y_1..y_r(B_1..B_s)),$$

де A_i – атомарна реляційна формула для представлення входження індивідних змінних $z_1..z_d$ у d-арне відношення P вигляду $P_{z_1...z_d}$; B_i – або атомарна реляційна формула, або рівність $x = y$, де x і y – індивідні змінні. Зокрема, вбудовані БЗЗ (ВБЗЗ) можна задати у вигляді

$$(\forall ab_1b_2c_1c_2d_1d_2)((Pab_1c_1d_1 \wedge Pab_2c_2d_2) \Rightarrow \exists d_3Pab_1c_2d_3).$$

Оскільки задача перевірки імплікації для таких видів обмежень, як ВБЗЗ, є нерозв'язною [29, 80], то неможливо побудувати алгоритм синтезу реляційної схеми у DK/NF.

Логічні зв'язки, встановлені на основі аналізу першоджерел та власних результатів між означеннями класичних НФ та основних некласичних НФ, відображені на діаграмі (див. рис. 1.2). Окрім розглянутих у даній роботі на діаграмі вказані логічні зв'язки між класичними НФ, які відображаються імплікаціями: НФБК$\Rightarrow$3НФ$\Rightarrow$2НФ$\Rightarrow$1НФ [62]. Означення розглянутих у роботі НФ зібрані у таблиці (додаток А).

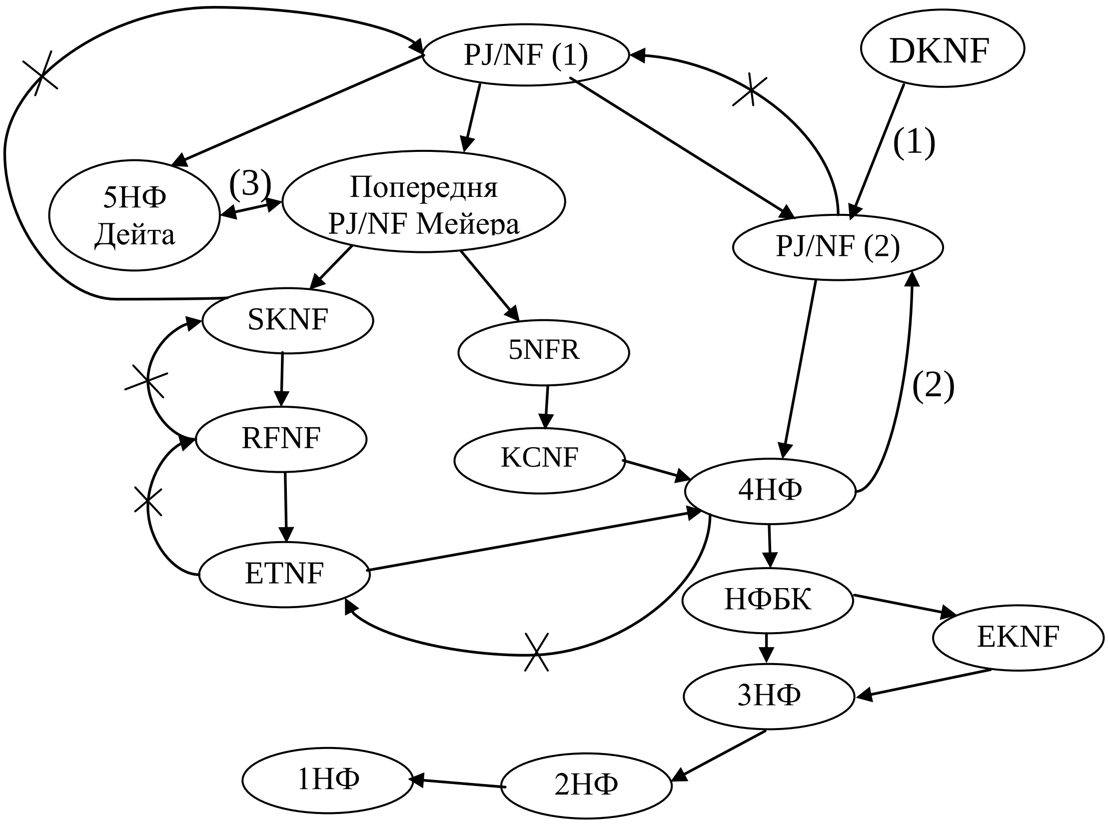

Рис. 1.2. Логічні зв'язки між означеннями НФ.

Умовні позначення:

(1) – за умови нескінченності універсального домена [26].

(2) – за умови, що усі ЗЗ є БЗЗ.

(3) – якщо вважати, що ЗЗ в означенні Дейта синтаксично слідують з множини Σ.

Окремим видом є шоста НФ (6НФ), запропонована для хронологічних БД у 2002 р.: «Відношення знаходиться у 6НФ тоді і тільки тоді, коли воно задовольняє усім нетривіальним залежностям з'єднання» [22]. З наведеного означення очевидно, що якщо відношення знаходиться у 6НФ, то воно знаходиться у PJ/NF за означенням 2).

Основні етапи розвитку абстрактній теорії залежностей та НФ в хронологічному порядку представлені в таблиці (додаток Б). В описі головних результатів робіт, по-можливості, зберігається термінологія їх авторів.

На сьогодні розвиток класичної теорії нормалізації реляційних БД відбувається у декількох напрямках, одним з яких є поширення її принципів на нечіткі реляційні БД (НРБД) [3, 15]. Для НРБД розглядаються концепції нечіткої ФЗ та неповної нечіткої ФЗ, які використовуються для визначення понять нечіткого ключа, транзитивного замикання та нечітких НФ, починаючи з 1НФ (Fuzzy Fit NF) і закінчуючи нечіткою НФБК (FBCNF). Зокрема, у роботі [3] наводяться алгоритми для побудови транзитивного замикання множини атрибутів, визначення нечітких потенційних ключів, зведення до нечіткої 2НФ, зведення до нечіткої 3НФ шляхом декомпозиції без втрат із збереженням множини нечітких ФЗ; в кінці роботи наводиться список літератури з даного напряму.

1.4 Основні результати розділу

Перший розділ присвячений огляду існуючої літератури з теорії нормалізації в реляційних (табличних) БД. По кожному з 49 джерел наведено головні результати. Основні результати першого розділу наведені в табл. 1.1.

Таблиця 1.1

Основні результати розділу 1

№ з.п.	Твердження	Стисле формулювання	Інтерпретація
(1)	(2)	(3)	(4)
1.	Твердження 1.1.	PJ/NF за означенням 1) $\Rightarrow$ PJ/NF за означенням 2). Обернена імплікація в загальному випадку не виконується.	Означення 1) для PJ/NF є більш сильним, ніж означення 2).

(1)	(2)	(3)	(4)
2.	Рис. 1.2.		Логічні зв’язки між означеннями НФ.

Аналіз наведеної літератури дозволяє зробити наступні висновки:

1) теорія нормалізації в реляційних БД, незважаючи на вагомі результати, носить фрагментарний характер;
2) накопичених теоретичних досліджень недостатньо для задоволення потреб розробників БД;
3) відсутність ефективного алгоритму зведення до PJ/NF (за означенням 2) пояснюється неможливістю побудови скінченної повної та коректної аксіоматики для ЗЗ;
4) потреби в автоматизації процесу нормалізації призводять до пошуку НФ сильніших ніж 4НФ але слабкіших ніж PJ/NF (за означенням 2).

РОЗДІЛ 2
АКСІОМАТИКА ФУНКЦІОНАЛЬНИХ ЗАЛЕЖНОСТЕЙ АРМСТРОНГА

2.1 Повнота аксіоматики Армстронга

Огляд наукової та навчальної літератури показав, що в ній відсутнє обґрунтування повноти аксіоматики Армстронга щодо ФЗ в реляційних БД даних, яке б задовольняло стандартним вимогам строгості та повноти математичного доведення.

Розглядаємо дві множини: A – множину атрибутів і D – універсальний домен, множина, з якої атрибути приймають значення в інтерпретаціях. Довільну (скінченну) множину атрибутів $R \subseteq A$ назвемо схемою [76]. В подальшому розгляді множини R і D зафіксовані.

Рядком схеми R називається іменна множина на парі R, D, проекція якої за першою компонентою рівна R (тобто по суті розглядається функція вигляду $s: R \to D$) [76].

Під таблицею t схеми R (позначатимемо $t(R)$) розуміємо множину рядків вказаної схеми R.

Скажемо, що на таблиці t виконується ФЗ (див., наприклад, [76, с. 71]) $X \to Y$, якщо для двох довільних рядків s_1, s_2 таблиці t, які збігаються на множині атрибутів X, має місце їх рівність і на множині атрибутів Y, тобто:

$$(X \to Y)(t) = true \overset{def}{\Leftrightarrow} \forall s_1, s_2 \in t\left(s_1|X = s_2|X \Rightarrow s_1|Y = s_2|Y\right).$$

Отже, з семантичної точки зору ФЗ – це предикат, заданий двома (скінченними) множинами атрибутів.

Скажемо, що таблиця t схеми R є моделлю множини ФЗ *F*, якщо кожна ФЗ $X \to Y \in F$ виконується на таблиці t:

$$t \text{ } \textit{модель} \text{ } F \overset{def}{\Leftrightarrow} \forall(X \to Y)(X \to Y \in F \Rightarrow (X \to Y)(t) = true).$$

Отже, довільна таблиця є моделлю порожньої множини ФЗ.

ФЗ $X \to Y$ семантично слідує з множини ФЗ *F*, якщо на кожній таблиці $t(R)$, яка є моделлю множини ФЗ *F*, виконується також ФЗ $X \to Y$:

$$F \vDash X \to Y \overset{def}{\Leftrightarrow} \forall t(R)(t \text{ модель } F \Rightarrow (X \to Y)(t) = true).$$

Лема 2.1 (аксіома рефлексивності Армстронга).

$\forall t(X \to Y)(t) = true$, де $Y \subseteq X$.□

Доведення. Розглянемо рядки s_1, s_2 таблиці t, для яких виконується рівність $s_1|X = s_2|X$. Обмежимо обидві частини цієї рівності за множиною Y: $(s_1|X)|Y = (s_2|X)|Y$. За властивістю оператора обмеження ($(U|Y)|Z = U|(Y \cap Z)$ згідно з [76, с. 24; 63]) маємо $s_1|(X \cap Y) = s_2|(X \cap Y)$. Звідси і з умови $Y \subseteq X$ випливає $s_1|Y = s_2|Y$, а отже $(X \to Y)(t) = true$.□

Наслідок 2.1. $\varnothing \vDash X \to Y$ для $Y \subseteq X$.□

ФЗ вигляду $X \to Y$, де $Y \subseteq X$, називається тривіальною.

Таким чином, тривіальні ФЗ виконуються на довільній таблиці; іншими словами, тривіальні ФЗ семантично слідують з порожньої множини ФЗ.

Лема 2.2 (правило поповнення Армстронга).

$(X \to Y)(t) = true \Rightarrow (X \cup Z \to Y \cup Z)(t) = true$ для $Z \subseteq R$.□

Доведення. Нехай на таблиці t виконується ФЗ $X \to Y$, покажемо, що на цій таблиці виконується і ФЗ $X \cup Z \to Y \cup Z$. Для цього розглянемо довільні рядки $s_1, s_2 \in t$, такі, що виконується $s_1 | (X \cup Z) = s_2 | (X \cup Z)$, звідси за аксіомою рефлексивності випливає: $s_1 | X = s_2 | X$ і $s_1 | Z = s_2 | Z$. Використовуючи властивість дистрибутивності обмежень відносно об'єднань [76, с. 24; 63], попередні рівності і імплікацію $s_1 | X = s_2 | X \Rightarrow s_1 | Y = s_2 | Y$, маємо ланцюжок рівностей $s_1 | (Y \cup Z) = s_1 | Y \cup s_1 | Z = s_2 | Y \cup$ $\cup s_2 | Z = s_2 | (Y \cup Z)$.□

Наслідок 2.2. $F \vDash X \to Y \Rightarrow F \vDash X \cup Z \to Y \cup Z$ для $Z \subseteq R$.□

Доведення. Нехай ФЗ $X \to Y$ семантично слідує з множини ФЗ *F*. Нехай далі таблиця t – довільна модель множини ФЗ *F*. Тоді за умовою $(X \to Y)(t) = true$. За попередньою лемою 2.2 (правило поповнення) $(X \bigcup Z \to Y \bigcup Z)(t) = true$. Отже, на довільній моделі множини ФЗ F виконується ФЗ $X \bigcup Z \to Y \bigcup Z$. Значить, за означенням семантичного слідування $F \vDash X \bigcup Z \to Y \bigcup Z$.□

Лема 2.3 (правило транзитивності Армстронга).

$$(X \to Y)(t) = true \wedge (Y \to Z)(t) = true \Rightarrow (X \to Z)(t) = true.\square$$

Доведення. Нехай на таблиці t виконуються ФЗ $X \to Y$ і $Y \to Z$; покажемо, що тоді на цій самій таблиці t виконується і ФЗ $X \to Z$. Для цього розглянемо довільні рядки $s_1, s_2 \in t$, такі, що $s_1 \mid X = s_2 \mid X$. Тоді, очевидно, що $s_1 \mid Y = s_2 \mid Y$. З останньої рівності випливає шукана рівність $s_1 \Big| Z = s_2 \Big| Z$.□

Наслідок 2.3. $F \vDash X \to Y \wedge F \vDash Y \to Z \Rightarrow F \vDash X \to Z$.□

Доведення. Припустимо, що ФЗ $X \to Y$, $Y \to Z$ семантично слідують з множини ФЗ *F*. Нехай далі таблиця t – довільна модель множини ФЗ F. Тоді за умовою $(X \to Y)(t) = true$ і $(Y \to Z)(t) = true$. Звідси за правилом транзитивності (лема 2.3) маємо $(X \to Z)(t) = true$. Отже, ФЗ $X \to Z$ виконується на довільній моделі множини ФЗ *F*; тобто $F \vDash X \to Z$.□

Правила поповнення і транзитивності назвемо правилами виведення; вони мають синтаксичну природу.

Скажемо, що ФЗ $X \to Y$ синтаксично слідує з множини ФЗ F ($F \vdash X \to Y$), якщо існує скінченна послідовність ФЗ $\varphi_1, \varphi_2, \ldots, \varphi_{m-1}, \varphi_m$, така, що $\varphi_m = X \to Y$ і $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить F, або отримана за яким-небудь правилом виведення (поповнення, транзитивності) з попередніх у цій послідовності ФЗ φ_j, φ_k, $j, k < i$ [71, с. 65]. Послідовність $\varphi_1, \varphi_2, \ldots, \varphi_{m-1}, \varphi_m$ назвемо доведенням, наслідуючи традиції математичної логіки [71].

Нехай задана деяка множина ФЗ F. Замикання $[F]_{\vdash}$ – це множина усіх ФЗ, які синтаксично слідують з F:

$$[F]_{\vdash} \overset{def}{=} \{X \to Y \mid F \vdash X \to Y\}.$$

Для спрощення позначень далі будемо записувати $[F]$ замість $[F]_{\vdash}$.

В наступній лемі викладено деякі природні властивості замикання.

Лема 2.4. Виконуються властивості:

1) $F \subseteq [F]$ (зростання);

2) $[[F]] = [F]$ (ідемпотентність);

3) $F \subseteq G \Rightarrow [F] \subseteq [G]$ (монотонність).□

Доведення. ■ Доведемо перше твердження. Нехай ФЗ $X \to Y \in F$, тоді $F \vdash X \to Y$ оскільки послідовність одиничної довжини $\langle X \to Y \rangle$ і є доведенням ФЗ $X \to Y$. ▪

■ Доведемо друге твердження. За властивістю 1) маємо: $[F] \subseteq [[F]]$. Доведемо обернене включення $[[F]] \subseteq [F]$. Нехай $X \to Y$ – довільна ФЗ, така, що $X \to Y \in [[F]]$. Тоді існує доведення $\varphi_1, \varphi_2, \ldots, \varphi_{m-1}, \varphi_m$, де $\varphi_m = X \to Y$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить $[F]$, або отримана за яким-небудь правилом виведення з попередніх у цій послідовності ФЗ φ_j, φ_k, $j, k < i$. Утворимо нову послідовність за такими правилами:

– якщо φ_i є аксіома рефлексивності, то запишемо цю ФЗ без змін;

– якщо $\varphi_i \in [F]$, то за означенням замикання ця ФЗ має скінченне доведення $\mu_1, \mu_2, \ldots, \mu_p$ з F. Замість ФЗ φ_i вставимо її доведення;

– якщо φ_i отримана за яким-небудь правилом виведення з попередніх у цій послідовності ФЗ φ_j, φ_k, $j, k < i$, то запишемо ФЗ φ_i без змін.

Очевидно, побудована таким чином послідовність є доведенням ФЗ $X \to Y$ з F, тобто $F \vdash X \to Y$, а отже $X \to Y \in [F]$. ▪

Ці властивості замикання наведені в роботі [73, с. 57].

■ Доведемо третє твердження. Нехай $X \to Y$ – довільна ФЗ, така, що $X \to Y \in [F]$. Тоді існує доведення $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$, де $\varphi_m = X \to Y$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить F, або отримана за яким-небудь правилом виведення з попередніх у цій послідовності ФЗ φ_j, φ_k, $j,k < i$. Оскільки множина ФЗ F є підмножиною множини ФЗ G, то, очевидно, що дане доведення є доведенням для ФЗ $X \to Y$, виходячи з множини ФЗ G: $X \to Y \in [G]$. ▪□

Таким чином, за термінологією [77] оператор $F \mapsto [F]$ є оператором замикання.

Нехай задана деяка множина ФЗ F, і нехай $\Im_F$ – сімейство усіх множин ФЗ G, які містять F, таких, що при застосуванні до них аксіом та правил виведення Армстронга не можна отримати жодної функціональної залежності, яка б не належала G:

$$\Im_F \overset{def}{=} \{G \mid F \subseteq G \wedge [G] \subseteq G\}.$$

Очевидно, що сім'я $\Im_F$ непорожня, вона містить, наприклад, множину всіх ФЗ $\{X \to Y \mid X \subseteq R \wedge Y \subseteq R\}$.

Лема 2.5. Сім'я $\Im_F$ замкнена відносно довільних перетинів. □

Доведення. Нехай $\Im'_F$ – деяка підсім'я сім'ї $\Im_F$ ($\Im'_F \subseteq \Im_F$); покажемо, що $G^* \overset{def}{=} \bigcap_{G \in \Im'_F} G \in \Im_F$. Для цього перевіримо дві умови на множини ФЗ, що є елементами сім'ї $\Im_F$.

■ Покажемо, по-перше, що $F \subseteq G^*$. Дійсно $F \subseteq G$ для всіх $G \in \Im'_F$, тому очевидно, що $F \subseteq \bigcap_{G \in \Im'_F} G = G^*$. ▪

■ Покажемо, по-друге, що $[G^*] \subseteq G^*$. Для цього розглянемо довільну ФЗ $X \to Y$, таку, що $X \to Y \in [G^*]$, і перевіримо, що виконується належність $X \to Y \in G^*$. Оскільки $X \to Y \in [G^*]$, то існує відповідне доведення (послідовність ФЗ). Шукане твердження доведемо індукцією по довжині доведення, де довжина доведення – кількість його елементів.

Базис індукції. Довжина доведення дорівнює одиниці. Тобто ФЗ $X \to Y$ є тривіальною або $X \to Y \in G^*$. В обох випадках $X \to Y \in G^*$ (використали той очевидний факт, що кожна множина G, $G \in \Im_F'$ містить всі тривіальні залежності, отже, і перетин $G^* = \bigcap_{G \in \Im_F'} G$ містить всі тривіальні залежності. ▪

Індуктивний крок. Нехай $\varphi_1, \varphi_2, ..., \varphi_m$, $m \geq 2$, – доведення ФЗ $X \to Y$, виходячи з множини ФЗ G^*. Розглянемо всі можливі випадки для $\varphi_m = X \to Y$.

Випадок, коли ФЗ $X \to Y$ тривіальна або $X \to Y \in G^*$, розглядається аналогічно як в базисі індукції. ▪

Нехай ФЗ $X \to Y$ отримується з ФЗ φ_i, φ_j, де $1 < i, j < m$ за правилом транзитивності. Очевидно, що послідовності $\varphi_1, ..., \varphi_i$, $\varphi_1, ..., \varphi_j$ є доведеннями довжини відповідно i та j, де $i, j < m$. Очевидно, що $\varphi_i, \varphi_j \in [G^*]$, причому довжини доведень менші ніж m. Згідно з індуктивним припущенням $\varphi_i, \varphi_j \in G^*$. Отже, $\varphi_i, \varphi_j \in G$ для всіх $G \in \Im_F'$; звідси за означенням сім'ї $\Im_F$, ФЗ $X \to Y \in G$ для всіх $G \in \Im_F$, тобто $X \to Y \in G^* = \bigcap_{G \in \Im_F'} G$.▪

Випадок, коли ФЗ $X \to Y$ отримується з попереднього елемента доведення за правилом поповнення, розглядається повністю аналогічно. ▪□

Нехай $F^* \overset{def}{=} \bigcap_{G \in \Im_F} G$ перетин всіх множин ФЗ з сім'ї $\Im_F$. З попередньої леми безпосередньо випливає наступний наслідок.

Наслідок 2.4. $F^* \in \Im_F$.□

Зауважимо, що сім'я ФЗ $\Im_F$ по суті є муровською сім'єю в розумінні абстрактної теорії решіток (див., наприклад, [47, с. 148]).

Лема 2.6. Виконується рівність: $[F] = F^*$.□

Доведення. ■ Покажемо спочатку включення $F^* \subseteq [F]$. Дійсно, оскільки $F \subseteq [F]$ та $[[F]] = [F]$ згідно з першими двома пунктами леми 2.4, то множина $[F]$ належить сім'ї $\Im_F$: $[F] \in \Im_F$.▪

■ Покажемо тепер обернене включення $[F] \subseteq F^*$. Дійсно, для довільної множини ФЗ $G \in \Im_F$ маємо включення $F \subseteq G$, звідси за пунктом 3 леми 2.4 маємо включення $[F] \subseteq [G]$; оскільки $[G] \subseteq G$, то $[F] \subseteq G$. Оскільки останнє включення виконується для всіх $G \in \Im_F$, то $[F] \subseteq F^*$.▪□

Висновок. Замикання $[F]$ – це найменша (в розумінні включення $\subseteq$) множина, що містить F, така, що при застосуванні до неї аксіом та правил виведення Армстронга не можна отримати жодної функціональної залежності, яка б не належала $[F]$.

З описаних вище аксіом і правил виведення можна отримати інші правила виведення (для спрощення практичного обчислення замикання $[F]$ множини ФЗ F).

Лема 2.7 (правило композиції).

$$\{X_1 \to Y_1,\ X_2 \to Y_2\} \vdash X_1 \cup X_2 \to Y_1 \cup Y_2 .\square$$

Доведення. Побудуємо доведення для ФЗ $X_1 \cup X_2 \to Y_1 \cup Y_2$:

1. $X_1 \to Y_1$;
2. $X_1 \cup X_2 \to Y_1 \cup X_2$ (з 1 за правилом поповнення);
3. $X_2 \to Y_2$;
4. $Y_1 \cup X_2 \to Y_1 \cup Y_2$ (з 3 за правилом поповнення);
5. $X_1 \cup X_2 \to Y_1 \cup Y_2$ (з 2 і 4 за правилом транзитивності). □

Звідси безпосередньо випливає наступний наслідок, якщо скористатися ідемпотентністю теоретико-множинного об'єднання.

Наслідок 2.5. $\{X \to Y_1,\ X \to Y_2\} \vdash X \to Y_1 \cup Y_2$.□

Наслідок 2.6. Для $n = 2,\ 3, \ldots$ має місце:

$$\{X \to Y_1, \ldots,\ X \to Y_n\} \vdash X \to Y_1 \cup \ldots \cup Y_n .\square$$

Доведення проводиться індукцією по n .□

Лема 2.8 (правило декомпозиції).

$$X \to Y_1 \cup Y_2 \vdash X \to Y_1,$$

$$X \to Y_1 \cup Y_2 \vdash X \to Y_2. \square$$

Доведення. Побудуємо доведення для ФЗ $X \to Y_1$:

1. $X \to Y_1 \cup Y_2$;
2. $Y_1 \cup Y_2 \to Y_1$ (аксіома рефлексивності);
3. $X \to Y_1$ (з 1 і 2 за правилом транзитивності).

Аналогічно будується доведення для ФЗ $X \to Y_2$. □

Правила композиції і декомпозиції наведені в [67].

Замиканням $[X]_F$ множини X (відносно множини ФЗ F) називається об'єднання правих частин всіх ФЗ вигляду $X \to Y$, які слідують (синтаксично) з множини F:

$$[X]_F \overset{def}{=} \bigcup_{X \to Y \in [F]} Y.$$

Для спрощення позначень далі параметр F явно вказувати не будемо.

В наступній лемі викладено властивості замикання множини X .

Лема 2.9. Виконуються властивості:

1) $X \subseteq [X]$;

2) $F \vdash X \to [X]$;

3) $X \to Z \notin [F] \Rightarrow Z \not\subseteq [X] \wedge [X] \subset R$. □

Доведення. ■ Перший пункт випливає з того факту, що тривіальні ФЗ містяться в замиканні довільної множини ФЗ, тобто $X \to X \in [F]$. ▪

■ Доведемо другий пункт. Замикання $[X]$ є скінченною множиною, нехай, для визначеності $[X] \overset{def}{=} \{A_1,\ldots,A_n\}$, де A_i – атрибути. Тоді, згідно з означенням замикання $[X]$, для кожного $i = 1, \ldots, n$ існує схема Y_i, така, що $A_i \in Y_i$ та $F \vdash X \to Y_i$. Згідно з правилом декомпозиції (лема 2.8) для кожного $i = 1, \ldots, n$ маємо $X \to Y_i \vdash X \to \{A_i\}$. Таким чином, маємо твердження $F \vdash X \to \{A_1\}$, ... ,

$F \vdash X \to \{A_n\}$. Залишається застосувати наслідок 2.6, згідно з яким $\{X \to \{A_1\}, \ldots, X \to \{A_n\}\} \vdash X \to [X]$. ▪

■ Доведемо останній пункт від супротивного. Нехай $X \to Z \notin [F]$, але заключення імплікації хибне. Тут можливі два випадки:

1) $[X] = R$. Оскільки за доведеним другим пунктом $F \vdash X \to [X]$, а значить $[X] \to Z$ тривіальна (нагадаємо, що $[X] = R$, а $Z \subseteq R$), то $F \vdash X \to Z$, тобто $X \to Z \in [F]$, прийшли до протиріччя; ▪

2) $Z \subseteq [X]$. Цей випадок розглядається аналогічно попередньому: $F \vdash X \to [X]$, $[X] \to Z$ тривіальна, звідки $F \vdash X \to Z$, що суперечить припущенню. ▪□

Твердження 2.1 (коректність аксіоматики Армстронга)**.** Якщо ФЗ $X \to Y$ синтаксично виводиться з множини ФЗ F, то ФЗ $X \to Y$ виводиться з F семантично:

$$F \vdash X \to Y \Rightarrow F \vDash X \to Y. \ \square$$

Доведення проводиться індукцією по довжині доведення.

Базис індукції. Довжина доведення дорівнює одиниці. Тобто ФЗ $X \to Y$ є або тривіальною або $X \to Y \in F$. В обох випадках (для першого треба скористатися наслідком 2.1) виконується $F \vDash X \to Y$. ▪

Індуктивний крок. Нехай $\varphi_1, \ldots, \varphi_m$, $m \geq 2$, доведення ФЗ $X \to Y$, виходячи з множини ФЗ F. Розглянемо всі можливі випадки для останнього елемента доведення $\varphi_m = X \to Y$.

Випадок, коли ФЗ $X \to Y$ тривіальна або належить множині ФЗ F, розглядається повністю аналогічно як в базисі індукції. ▪

Нехай ФЗ φ_m отримується з ФЗ φ_i, де $i < m$, за правилом поповнення. Очевидно, що $F \vdash \varphi_i$; за індуктивним припущенням маємо $F \vDash \varphi_i$. Залишається скористатися наслідком 2.2. ▪

Випадок, коли ФЗ φ_m отримується з попередніх ФЗ доведення за правилом транзитивності, розглядається аналогічно з заміною наслідку 2.2 на наслідок 2.3. ▪□

Зауважимо, що логічна схема доведення попереднього твердження та сама, що і для доведення леми 2.5 (точніше кажучи, доведення включення $[G^*] \subseteq G^*$).

Твердження 2.2 (повнота аксіоматики Армстронга). Якщо ФЗ $X \to Y$ семантично виводиться з множини ФЗ F, то ФЗ $X \to Y$ синтаксично виводиться з F:

$$F \vDash X \to Y \Rightarrow F \vdash X \to Y. \ \square$$

Доведення проводиться від супротивного. Отже, нехай множина ФЗ F та ФЗ $X \to Y$ такі, що $F \vDash X \to Y$, але не виконується $F \vdash X \to Y$, тобто $X \to Y \notin [F]$.

Прийдемо до протиріччя (з $F \vDash X \to Y$), побудувавши таку модель множини ФЗ F, що на ній ФЗ $X \to Y$ не буде виконуватися.

Зафіксуємо в універсальному домені два різних елементи a та b. Нехай схема R складається з атрибутів $A_1, \ldots, A_n$. Таблиця t схеми R містить рядки s_1 і s_2, які задані наступним чином:

$$s_1(A_1) = \ldots = s_1(A_n) = a,$$

$$s_2(A_i) = \begin{cases} a, & \text{якщо} \ \ A_i \in [X], \\ b, & \text{якщо} \ \ A_i \notin [X], \end{cases}$$

де $i = 1, \ldots, n$.

З третього пункту леми 2.9 випливає, що $[X] \subset R$; значить, рядки s_1, s_2 розрізняються на всіх атрибутах з $R \setminus [X]$ (див. рис. 2.1).

	$[X]$			$R \setminus [X]$		
	A_1	...	A_{i_m}	$A_{i_{m+1}}$	…	A_n
s_1	a	...	a	a	…	a
s_2	a	...	a	b	...	b

Рис. 2.1. Таблиця $t \stackrel{def}{=} \{s_1, s_2\}$ з доведення твердження 2.2.

Згідно з пунктом 1 леми 2.9 $X \subseteq [X]$, а згідно з пунктом 3 – $Y \not\subseteq [X] \subset R$, тобто $Y \cap (R \setminus [X]) \neq \varnothing$. З побудови рядків s_1, s_2 випливає, що $s_1 | X = s_2 | X$, але $s_1 | Y \neq s_2 | Y$, таким чином $(X \to Y)(t) = false$. ▪

■ Залишається показати, що таблиця t є моделлю множини ФЗ F. Для цього розглянемо довільну ФЗ $W \to Z$ з F та покажемо, що $(W \to Z)(t) = true$, тобто імплікація $s_1 | W = s_2 | W \Rightarrow s_1 | Z = s_2 | Z$ виконується.

Для множини атрибутів W можливі два випадки.

По-перше, нехай $W \cap (R \setminus [X]) \neq \varnothing$, тоді $s_1 | W \neq s_2 | W$ і імплікація тривіально істинна.

По-друге, нехай $W \cap (R \setminus [X]) = \varnothing$, тобто $W \subseteq [X]$; тоді (нагадаємо, що $X \subseteq [X]$ згідно з першим пунктом леми 2.9) $s_1 | W = s_2 | W$, тому треба перевірити рівність $s_1 | Z = s_2 | Z$. Цю рівність ми доведемо, довівши включення $Z \subseteq [X]$. Для цього розглянемо наступне доведення, виходячи з множини ФЗ F:

1. Доведення ФЗ $X \to [X]$ з F (нагадаємо, що згідно пункту 2 леми 2.9 $F \vdash X \to [X]$);

2. $[X] \to W$ (аксіома рефлексивності; нагадаємо, що розглядається випадок $W \subseteq [X]$);

3. $X \to W$ (правило транзитивності застосовано до ФЗ $X \to [X]$, яка є останнім елементом доведення з пункту 1, та ФЗ $[X] \to W$ з пункту 2);

4. $W \to Z$ (елемент множини F);

5. $X \to Z$ (правило транзитивності застосовано до ФЗ з пунктів 3 та 4).

Отже, маємо $F \vdash X \to Z$, тобто $Z \subseteq [X]$; звідси випливає, що $s_1 | Z = s_2 | Z$ за побудовою рядків s_1 і s_2. ▪□

Твердження про повноту аксіоматики Армстронга наводиться в [с. 58, 73].

Теорема 2.1. Відношення семантичного та синтаксичного слідування збігаються:

$$F \vDash X \to Y \Leftrightarrow F \vdash X \to Y. \square$$

Зробимо декілька зауважень про отриманий побічний результат. Виявляється, що оператор $X \mapsto [X]$ є теж оператором замикання.

Твердження 2.3. Виконуються властивості:

1) $X \subseteq Y \Rightarrow [X] \subseteq [Y]$ (монотонність);

2) $[[X]] = [X]$ (ідемпотентність).□

Доведення. ■ Доведемо першу властивість. Покажемо включення $[X] \subseteq [Y]$, побудувавши доведення для ФЗ $Y \to [X]$ (виходячи з множини ФЗ F):

1. $Y \to X$ (аксіома рефлексивності; нагадаємо, що $X \subseteq Y$);

2. Доведення ФЗ $X \to [X]$ з F (згідно пункту 2 леми 2.9 $F \vdash X \to [X]$);

3. $Y \to [X]$ (правило транзитивності застосовано до ФЗ $Y \to X$ з пункту 1 та ФЗ $X \to [X]$, яка є останнім елементом доведення з пункту 2).

Отже, $F \vdash Y \to [X]$, а це означає, що множина атрибутів $[X]$ включається в замикання $[Y]$ множини атрибутів Y, тобто $[X] \subseteq [Y]$. ▪

■ Доведемо другу властивість. Включення $[X] \subseteq [[X]]$ випливає з пункту 1 леми 2.9, тому доведемо обернене включення $[[X]] \subseteq [X]$. Нехай атрибут $A \in [[X]]$; покажемо, що $A \in [X]$. Для цього побудуємо доведення для ФЗ $X \to \{A\}$ (виходячи з множини ФЗ F):

1. Доведення ФЗ $[X] \to \{A\}$ з F ($F \vdash [X] \to \{A\}$ за означенням замикання $[[X]]$ множини $[X]$);

2. Доведення ФЗ $X \to [X]$ з F (згідно пункту 2 леми 2.9 $F \vdash X \to [X]$);

3. $X \to \{A\}$ (правило транзитивності застосовано до ФЗ $X \to [X]$, яка є останнім елементом доведення з пункту 2, та ФЗ $[X] \to \{A\}$, яка також є останнім елементом доведення з пункту 1).

Отже, $F \vdash X \to \{A\}$; за означенням замикання $[X]$ множини X випливає, що $A \in [X]$. ▪□

2.2 Критерій повноти аксіоматики Армстронга в термінах потужностей множини атрибутів та універсального домена

Детальний аналіз наведеного доведення основного результату про збіжність відношень синтаксичного ($\vdash$) та семантичного ($\vDash$) слідування показує, що воно проведено в припущенні: $|D|\geq 2$ та $|R|\geq 2$, тобто універсальний домен містить не менше 2 елементів та схема R має щонайменше 2 атрибути. Для повноти викладення треба розглянути і решту випадків, коли $|D|\leq 1$ або $|R|\leq 1$.

Лема 2.10. Всі ФЗ виконуються на порожній таблиці. ФЗ вигляду $X\to\varnothing$, зокрема, ФЗ вигляду $\varnothing\to\varnothing$ виконується на довільній таблиці. □

Доведення випливає безпосередньо з означень. □

Лема 2.11. Довільна таблиця є моделлю порожньої множини ФЗ. □

Доведення випливає безпосередньо з означень. □

Лема 2.12. Множина тривіальних ФЗ замкнена відносно правил поповнення та транзитивності. □

Доведення. Імплікація $X\to Y$ – тривіальна $\Rightarrow X\bigcup Z\to Y\bigcup Z$ – тривіальна є очевидною. Імплікація $X\to Y$, $Y\to Z$ – тривіальні $\Rightarrow X\to Z$ – тривіальна перевіряється безпосередньо.

Дійсно, нехай $Y\subseteq X$, тоді, очевидно, що $Y\bigcup Z\subseteq X\bigcup Z$.▫

Якщо $Y\subseteq X$, а $Z\subseteq Y$, то, тривіально, що $Z\subseteq X$.▫□

Лема 2.13. Всі ФЗ виконуються на довільній однорядковій таблиці. □

Доведення випливає безпосередньо з означень. □

Нехай F – множина ФЗ; через $(F)_{ntr}$ (None TRivial) позначимо підмножину множини F, що отримується вилученням тривіальних ФЗ.

Лема 2.14. Виконуються наступні твердження:

1) $F\vdash\varphi$, якщо $\varphi\in F$;

2) $F\vDash\varphi$, якщо $\varphi\in F$;

3) $F\vdash\varphi\Leftrightarrow(F)_{ntr}\vdash\varphi$;

4) $F \vDash \varphi \Leftrightarrow (F)_{ntr} \vDash \varphi$;

5) $F \vdash \varphi$, якщо φ – тривіальна ФЗ;

6) $F \vDash \varphi$, якщо φ – тривіальна ФЗ. □

Доведення випливає безпосередньо з означень відношень синтаксичного та семантичного слідувань. □

Залежність збіжності відношень синтаксичного та семантичного слідувань при різних значеннях потужностей множин R та D вказана у табл. 2.1. Дана таблиця містить три рядки та три стовпчики. Комірки будемо позначати парами, перша компонента – номер рядка, друга – номер стовпчика. “+” (відповідно “–”) в комірці означає, що при вказаних припущеннях відношення $\vdash$ і $\vDash$ збігаються (не збігаються).

Таблиця 2.1

Всі варіанти потужностей множин R та D

		1	2	3
	D \ R	$\|R\|=0$	$\|R\|=1$	$\|R\|\geq 2$
1	$\|D\|=0$	+	–	–
2	$\|D\|=1$	+	–	–
3	$\|D\|\geq 2$	+	+	+

Твердження 2.4. Заповнення табл. 2.1 коректне. □

Доведення проводиться розглядом всіх можливих 9 випадків.

■ Випадок (1,1): $|D| = |R| = 0$. Тоді $t_{\varnothing}$, t_{ε} – всі можливі таблиці в інтерпретації; ці таблиці можуть бути моделями тільки таких множин ФЗ: $F = \varnothing$ або $F = \{\varnothing \to \varnothing\}$, причому ФЗ вигляду $\varnothing \to \varnothing$ є тривіальною.

Множина всіх ФЗ в інтерпретації задається з урахуванням очевидної рівності $I_R = \{X \to Y \mid X \in 2^R \ \wedge \ Y \in 2^R\}$. У випадку, що розглядається ($R = \varnothing$), $I_{\varnothing} = \{\varnothing \to \varnothing\}$, оскільки $2^{\varnothing} = \{\varnothing\}$.

Крім того, множина ФЗ при розгляді моделей може бути і порожньою.

Покажемо, що відношення $\vdash$ і $\vDash$ співпадають; для цього розглянемо всі можливі випадки:

$$\varnothing \vdash \varnothing \to \varnothing \Leftrightarrow \varnothing \vDash \varnothing \to \varnothing \text{ (для випадку } F = \varnothing);$$

$$\{\varnothing \to \varnothing\} \vdash \varnothing \to \varnothing \Leftrightarrow \{\varnothing \to \varnothing\} \vDash \varnothing \to \varnothing \text{ (для випадку } F = \{\varnothing \to \varnothing\}).$$

Ці еквівалентності випливають з того, що виконуються твердження: $\varnothing \vdash \varnothing \to \varnothing$ (пункт 5 леми 2.14), $\varnothing \vDash \varnothing \to \varnothing$ (пункт 6 леми 2.14), $\{\varnothing \to \varnothing\} \vdash \varnothing \to \varnothing$ (пункт 1 леми 2.14), $\{\varnothing \to \varnothing\} \vDash \varnothing \to \varnothing$ (пункт 2 леми 2.14). ▪

■ Випадок (2,1): $|D| = 1$, $R = \varnothing$. Тоді $t_{\varnothing}$, t_{ε} – всі можливі таблиці в інтерпретації. Розглядається аналогічно випадку (1,1). ▪

■ Випадок (3,1): $|D| \geq 2$, $R = \varnothing$. Розглядається аналогічно випадкам (1,1) і (2,1). ▪

■ Випадок (1,2): $|R| = 1$, $D = \varnothing$. $t_{\varnothing}$ – єдина можлива таблиця в інтерпретації. Покладемо для визначеності $R \overset{def}{=} \{A\}$, звідси можливими ФЗ є: $\varnothing \to \varnothing$, $\varnothing \to \{A\}$, $\{A\} \to \varnothing$, $\{A\} \to \{A\}$. Покажемо, що відношення $\vdash$ і $\vDash$ не співпадають; для цього наведемо такий контрприклад: $\varnothing \vdash \varnothing \to \{A\} \Leftrightarrow \varnothing \vDash \varnothing \to \{A\}$ (з припущення $F = \varnothing$).

Вказана еквівалентність не виконується тому, що $\varnothing \vDash \varnothing \to \{A\}$ має місце згідно з лемою 2.10, а $\varnothing \vdash \varnothing \to \{A\}$ не має місця згідно з лемою 2.12. Формальне доведення проводиться від супротивного. Дійсно, припустимо, що $\varnothing \vdash \varnothing \to \{A\}$ має місце, тоді ФЗ $\varnothing \to \{A\}$ повинна належати замиканню $[\varnothing]$. Згідно з пунктом 5 леми 2.14 з множини $\varnothing$ синтаксично слідують усі можливі тривіальні ФЗ, а саме: $\varnothing \to \varnothing$, $\{A\} \to \varnothing$, $\{A\} \to \{A\}$, а згідно з лемою 2.12 замикання $[\varnothing]$ містить лише тривіальні ФЗ. Оскільки ФЗ $\varnothing \to \{A\}$ не є тривіальною, то вона не належить замиканню, а отже, не має доведення з $\varnothing$. Прийшли до протиріччя з припущенням. ▪

■ Випадок (2,2): $|D| = 1$, $|R| = 1$. Нехай для визначеності $R \overset{def}{=} \{A\}$ та $D \overset{def}{=} \{d\}$. В інтерпретації можливі тільки таблиці: $t_{\varnothing}$ і $t_0 \overset{def}{=} \{\{\langle A, d \rangle\}\}$ (рис. 2.2).

У цьому випадку відношення $\vdash$ і $\vDash$ також не співпадають. Приклад той самий, що і в попередньому випадку (1,2): $\varnothing \vDash \varnothing \to \{A\}$ виконується згідно з лемами 2.10 (для випадку $t_\varnothing$) та 2.13 (для випадку t_0), а $\varnothing \vdash \varnothing \to \{A\}$ не виконується згідно з лемою 2.12. ▪

■ Випадок (3,2): $|D| \geq 2$, $|R| = 1$. Нехай для визначеності $R \stackrel{def}{=} \{A\}$. Всі можливі ФЗ як і для випадку (1,2): $\varnothing \to \varnothing$, $\varnothing \to \{A\}$, $\{A\} \to \varnothing$, $\{A\} \to \{A\}$. З урахуванням леми 2.14 (пункти 3 і 4) достатньо перевірити еквівалентність $\varnothing \vdash \varnothing \to \{A\} \Leftrightarrow \varnothing \vDash \varnothing \to \{A\}$. Дійсно, обидві частини еквівалентності хибні: $\varnothing \vdash \varnothing \to \{A\}$ хибне згідно з лемою 2.12 (доведення те ж саме, що і у випадку (1,2)), а $\varnothing \vDash \varnothing \to \{A\}$ хибне, оскільки ФЗ $\varnothing \to \{A\}$ не виконується на дворядковій таблиці $t_0' \stackrel{def}{=} \left\{\{\langle A, d_1 \rangle\}, \{\langle A, d_2 \rangle\}\right\}$ (рис. 2.2); вище покладено, що $d_1, d_2 \in D$ та різні. ▪

A
d

A
d_1
d_2

Рис. 2.2. Таблиця t_0 для випадку (2.2) (зліва),

таблиця t_0' для випадку (3.2) (справа).

■ Випадок (1,3): $D = \varnothing$, $|R| \geq 2$. $t_\varnothing$ – єдина таблиця в інтерпретації. Відношення $\vdash$ і $\vDash$ не співпадають. Приклад той самий, що і для випадку (1,2): $\varnothing \vdash \varnothing \to \{A\}$ не виконується, а $\varnothing \vDash \varnothing \to \{A\}$ виконується. ▪

■ Випадок (2,3): $|D| = 1$, $|R| \geq 2$. Відношення $\vdash$ і $\vDash$ не співпадають. Приклад той самий, що і для випадків (1,2), (2,2). ▪

■ Випадок (3,3): $|D| \geq 2$, $|R| \geq 2$. Відношення $\vdash$ і $\vDash$ співпадають. Саме про це говорить доповнене, реконструйоване доведення результату Армстронга з підрозділу 2.1. ▪□

Теорема 2.2 (критерій повноти аксіоматики Армстронга). Відношення $\vdash$ і $\models$ співпадають тоді і тільки тоді, коли в інтерпретації $|D| \geq 2$ або $|R| = 0$.□

Доведення випливає з заповнення табл. 2.1. □

2.3 Незалежність складових аксіоматики Армстронга

Нагадаємо, що аксіоматика Армстронга при зафіксованій множині атрибутів R складається з:

– аксіом рефлексивності $X \rightarrow Y$, $Y \subseteq X$;

– правил поповнення $\dfrac{X \rightarrow Y}{X \bigcup Z \rightarrow Y \bigcup Z}$ для $Z \subseteq R$;

– правил транзитивності $\dfrac{X \rightarrow Y, Y \rightarrow Z}{X \rightarrow Z}$.

Лема 2.15. Аксіоми рефлексивності є незалежними від правил виведення аксіоматики Армстронга. □

Доведення є тривіальним, оскільки інших аксіом в аксіоматиці Армстронга немає. □

Лема 2.16. Правила транзитивності є незалежними від аксіом рефлексивності і правил поповнення. □

Доведення. Для доведення достатньо вказати множину ФЗ F, для якої $[F]_{\vdash tr} \subset [F]_{\vdash}$, де $[F]_{\vdash tr}$ – замикання множини ФЗ F, побудоване за допомогою аксіом рефлексивності і правил поповнення, а $[F]_{\vdash}$ – замикання множини ФЗ F відносно всіх аксіом і правил виведення (див. підрозділ 2.1).

Зафіксуємо множини атрибутів X і Z такі, що $X \bigcap Z = \varnothing$, і розглянемо множину ФЗ F, яка задовольняє умовам $\forall Y \subseteq R (X \rightarrow Y \in F \wedge Y \rightarrow Z \in F \Rightarrow$ $\Rightarrow X \rightarrow Z \notin F)$. Покажемо, що ФЗ $X \rightarrow Z$ не може бути виведена лише із застосуванням аксіом рефлексивності та правил поповнення.

Оскільки $Z \bigcap X = \varnothing$, то $Z \not\subseteq X$, отже, ФЗ $X \rightarrow Z$ не є тривіальною. ▪

Припустимо, що ФЗ $X \to Z$ отримана з деякої ФЗ $U \to V$ за правилом поповнення, тобто існує така множина атрибутів $W \neq \varnothing$, що $U \bigcup W = X$ і $V \bigcup W = Z$. Тоді $Z \bigcap X = (V \bigcup W) \bigcap (U \bigcup W) \supseteq W \neq \varnothing$, що суперечить умові $Z \bigcap X = \varnothing$. Отримали протиріччя з припущенням, отже, $[F]_{\vdash tr} \subset [F]_{\vdash}$ і правило транзитивності є незалежним. □

Лема 2.17. Правила поповнення є незалежними від аксіом рефлексивності і правил транзитивності. □

Доведення. Розглянемо множину ФЗ $F = \{X \to \{A\} \mid A \notin X \wedge X \neq \varnothing\}$ і покажемо, що ФЗ вигляду $X \to X \bigcup \{A\}$, де $A \notin X$ і $X \neq \varnothing$, не можна вивести лише за допомогою аксіом рефлексивності і правил транзитивності.

Припустимо, що ФЗ $X \to X \bigcup \{A\} \in F$, тоді її права частина є одноелементною множиною. Останнє виконується, якщо:

– $A \in X$ і $X = \{A\}$;

– $A \notin X$ і $X = \varnothing$.

Оскільки кожен з зазначених випадків суперечить визначенню множини ФЗ F, то $X \bigcup \{A\}$ не є одноелементною множиною і ФЗ $X \to X \bigcup \{A\} \notin F$.

Оскільки $A \notin X$, то ФЗ $X \to X \bigcup \{A\}$ не є тривіальною.

Припустимо, ФЗ $X \to X \bigcup \{A\}$ отримана за правилами транзитивності. Враховуючи, що

1) множина тривіальних ФЗ замкнена відносно правила транзитивності (лема 2.12);

2) результатом застосування правила транзитивності до ФЗ з одноелементною правою частиною є ФЗ, права частина якої є знову одноелементною множиною;

3) результатом застосування правила транзитивності до тривіальної ФЗ і ФЗ з одноелементною правою частиною є ФЗ з одноелементною правою частиною;

4) результатом застосування правила транзитивності до ФЗ виду $X \to \{A\}$ і тривіальної ФЗ $\{A\} \to \{A\}$ є ФЗ виду $X \to \{A\}$,

приходимо до висновку, що ФЗ $X \to X \bigcup \{A\}$ повинна бути або тривіальною або її права частина є одноелементною множиною, що суперечить визначенню цієї ФЗ. Отримали протиріччя з припущенням, отже правило поповнення є незалежним. □

Теорема 2.3. Аксіоматика Армстронга є незалежною. □

Доведення випливає безпосередньо з лем 2.15, 2.16 і 2.17. □

2.4 Алгебра функціональних залежностей в табличних базах даних

Задамо алгебру $A = \langle \Psi, \Omega \rangle$, носієм якої є множина ФЗ $\Psi = \left\{ X \to Y \mid X, Y \in 2^R \right\}$, де 2^R – булеан множини R; а сигнатура Ω складається з операцій, які задаються наступним чином:

– нуль-арні операції $X \to Y^{(0)}$, де $X, Y \subseteq R$, причому $Y \subseteq X$, які відповідають аксіомам рефлексивності в аксіоматиці Армстронга;

– унарні параметричні операції $R_Z'^{(1)}$, які відповідають правилам поповнення в аксіоматиці Армстронга:

$$R_Z'^{(1)} : \Psi \to \Psi,\ Z \in 2^R,\ dom\ R_Z'^{(1)} = \{X \to Y \mid X, Y \in 2^R\},$$

$$R_Z'^{(1)}(X \to Y) = X \bigcup Z \to Y \bigcup Z\ ;$$

– бінарна часткова операція $R''^{(2)}$, яка відповідає правилу транзитивності в аксіоматиці Армстронга:

$$R''^{(2)} : \Psi \times \Psi \tilde{\to} \Psi,\ dom\ R''^{(2)} = \left\{ \langle X \to Y', Y'' \to Z \rangle \mid Y' = Y'' \right\},$$

$$R''^{(2)}(X \to Y, Y \to Z) = X \to Z\ .$$

Таким чином, алгебра ФЗ є частковою [72] і задається як:

$$A = \left\langle \left\{ X \to Y \right\}_{X, Y \in 2^R}, \{R_Z'^{(1)}, R''^{(2)}, X \to Y^{(0)}\}_{Z, X, Y \in 2^R, Y \subseteq X} \right\rangle.$$

Позначимо множину аксіом рефлексивності, які задаються нуль-арними операціями $X \to Y^{(0)}$ через $\Psi_r = \{X \to Y \mid X, Y \in 2^R \wedge Y \subseteq X\}$. Очевидно, що $A' = \left\langle \Psi_r, \{R_Z'^{(1)}, R''^{(2)}\}_{Z,X,Y \in 2^R, Y \subseteq X} \right\rangle$ є підалгеброю алгебри A, оскільки множина тривіальних ФЗ Ψ_r є замкненою відносно вказаних операцій.

Лема 2.18. Множина $\Psi_{r_0} = \{\{A\} \to \varnothing \mid A \in R\}$ породжує множину тривіальних залежностей Ψ_r за допомогою операцій $R_Z'^{(1)}$ і $R''^{(2)}$. □

Доведення. Покажемо, що довільну тривіальну ФЗ $X \to Y \in \Psi_r$, де $Y \subseteq X$ можна отримати з множини Ψ_{r_0} за допомогою операцій $R_Z'^{(1)}$ і $R''^{(2)}$.

Нехай схема R задається множиною атрибутів $\{A_1, A_2, \ldots, A_n\}$. Не втрачаючи загальності представимо множини атрибутів X і Y, де $Y \subseteq X$, відповідно $\{A_1, A_2, \ldots, A_m\}$ і $\{A_1, A_2, \ldots, A_k\}$, де $k \le m \le n$.

Якщо $k = m$, то тривіальну ФЗ $X \to Y$ можна отримати з ФЗ $A_i \to \varnothing$, де i – фіксоване, причому $i \le m$, виконавши операцію:

$$R'^{(1)}_{\{A_1,\ldots,A_m\}}(\{A_i\} \to \varnothing) = \{A_i\} \cup \{A_1, \ldots, A_m\} \to \varnothing \cup \{A_1, \ldots, A_m\},$$

результатом якої є ФЗ $\{A_1, \ldots, A_m\} \to \{A_1, \ldots, A_m\}$.

У випадку $k < m$ ФЗ $X \to Y$ можна отримати з множини ФЗ $\{\{A_{k+1}\} \to \varnothing, \{A_{k+2}\} \to \varnothing, \ldots, \{A_m\} \to \varnothing\}$, виконавши наступну послідовність операцій:

крок 1: $R'^{(1)}_{\{A_1,\ldots,A_k\}}(\{A_{k+1}\} \to \varnothing) = \{A_{k+1}\} \cup \{A_1, \ldots, A_k\} \to \varnothing \cup \{A_1, \ldots, A_k\}$,

результатом якої є ФЗ $\{A_1, \ldots, A_{k+1}\} \to \{A_1, \ldots, A_k\}$; якщо $k+1 = m$, то ФЗ $X \to Y$ отримана, інакше ($k+1 < m$) виконуємо крок 2, який складається з двох операцій:

1) $R'^{(1)}_{\{A_1,\ldots,A_{k+1}\}}(\{A_{k+2}\} \to \varnothing) = \{A_{k+2}\} \cup \{A_1, \ldots, A_{k+1}\} \to \varnothing \cup \{A_1, \ldots, A_{k+1}\}$,

результатом якої є ФЗ $\{A_1, \ldots, A_{k+2}\} \to \{A_1, \ldots, A_{k+1}\}$;

2) $R''^{(2)}(\{A_1,...,A_{k+2}\} \to \{A_1,...,A_{k+1}\}, \{A_1,...,A_{k+1}\} \to \{A_1,...,A_k\}) =$
$= \{A_1,...,A_{k+2}\} \to \{A_1,...,A_k\}$.

Перепишемо крок 2 для $k+i<m$ $(i=1,2,...,m-k-1)$:

1) $R'^{(1)}_{\{A_1,...,A_{k+i}\}}(\{A_{k+i+1}\} \to \varnothing) = \{A_{k+i+1}\} \bigcup \{A_1,...,A_{k+i}\} \to \varnothing \bigcup \{A_1,...,A_{k+i}\}$,

результатом якої є ФЗ $\{A_1,...,A_{k+i+1}\} \to \{A_1,...,A_{k+i}\}$;

2) $R''^{(2)}(\{A_1,...,A_{k+i+1}\} \to \{A_1,...,A_{k+i}\}, \{A_1,...,A_{k+i}\} \to \{A_1,...,A_k\}) =$
$= \{A_1,...,A_{k+i+1}\} \to \{A_1,...,A_k\}$.

Виконавши крок 2 $m-k-1$ разів отримаємо ФЗ $\{A_1,...,A_m\} \to \{A_1,...,A_k\}$, тобто, ФЗ $X \to Y$. □

Таким чином, замикання множини $\Psi_{r_0} \subseteq \Psi_r$ операціями збіднення алгебри A' ($R'^{(1)}_Z$ і $R''^{(2)}$ – операції збіднення, позначається: $[\Psi_{r_0}]_{\left\{R'^{(1)}_Z,\ R''^{(2)}\right\}}$) співпадає з множиною тривіальних залежностей Ψ_r, тобто:

$$[\Psi_{r_0}]_{\left\{R'^{(1)}_Z,\ R''^{(2)}\right\}} = \Psi_r .$$

В свою чергу, результат виконання операції $R'^{(1)}_Z$ можна представити за допомогою операції $R'^{(1)}_{\{A\}}$, $A \in R$, отже, маємо:

$$[\Psi_{r_0}]_{\left\{R'^{(1)}_{\{A\}},\ R''^{(2)}\right\}} = \Psi_r .$$

Лема 2.19. Множина Ψ_{r_0} – найменша множина, яка породжує множину Ψ_r за допомогою основних операцій $R'^{(1)}_{\{A\}}$, $A \in R$ і $R''^{(2)}$. □

Доведення. Якщо схема $R = \varnothing$, то очевидно, що $\Psi_{r_0} = \Psi_r = \{\varnothing \to \varnothing\}$.

Розглянемо випадок, коли схема $R \neq \varnothing$. Зафіксуємо деякий атрибут $B \in R$ і розглянемо множину $\Psi'_{r_0} \subset \Psi_{r_0}$ таку, що $\Psi'_{r_0} = \left\{\{A\} \to \varnothing \mid A \in R \setminus \{B\}\right\}$. Покажемо, що множина Ψ'_{r_0} не породжує множину $\Psi_{r_0} \subseteq \Psi_r$ за допомогою операцій $R'^{(1)}_{\{A\}}$, $A \in R$ і $R''^{(2)}$; для цього розглянемо можливі випадки для отримання ФЗ $\{B\} \to \varnothing$:

1) ФЗ $\{B\} \to \varnothing$ не може бути результатом виконання операції $R'^{(1)}_{\{A\}}$, $A \in R$, застосованої до деякої тривіальної ФЗ $\psi \in \Psi'_{r_0}$, оскільки містить порожню множину у своїй правій частині;

2) припустимо, що тривіальна ФЗ $\{B\} \to \varnothing$ є результатом виконання операції $R''^{(2)}$, застосованої до деяких тривіальних ФЗ $\varphi, \psi \in \Psi'_{r_0}$. Тоді повинна існувати така множина атрибутів Y, де $\varnothing \subseteq Y \subseteq \{B\}$, що виконується ФЗ $\{B\} \to Y$ і $Y \to \varnothing$. Випадки, коли $Y = \varnothing$ або $Y = \{B\}$ вимагають використання самої ФЗ $\{B\} \to \varnothing$:

a. $\varphi = \{B\} \to \varnothing$ і $\psi = \varnothing \to \varnothing$;

b. $\varphi = \{B\} \to \{B\}$ і $\psi = \{B\} \to \varnothing$.

Оскільки множини атрибутів Y, яка б задовольняла умові $\varnothing \subset Y \subset \{B\}$ не існує, приходимо до висновку, що множина Ψ'_{r_0} не породжує множину $\Psi_{r_0} \subseteq \Psi_r$, а отже, не породжує множину Ψ_r.□

Таким чином, алгебра ФЗ задається як своє збіднення наступного вигляду:

$$A = \left\langle \{X \to Y\}_{X,Y \in 2^R}, \{R'^{(1)}_{\{A\}}, R''^{(2)}, \{A\} \to \varnothing^{(0)}\}_{A \in R} \right\rangle.$$

Зауважимо, що замикання $[F]$ множини ФЗ F, де $F \subseteq \{X \to Y\}_{X,Y \in 2^R}$ можна розглядати як замикання множини $F \cup \Psi_{r_0}$ операціями $R'^{(1)}_{\{A\}}$, $A \in R$ і $R''^{(2)}$:

$$[F \cup \Psi_{r_0}]_{\left\{R'^{(1)}_{\{A\}},\, R''^{(2)}\right\}}.$$

2.5 Основні результати розділу

У другому розділі розглядається аксіоматика Армстронга для ФЗ.

Введено поняття семантичного та синтаксичного слідувань. Розглянуті властивості замикання множини атрибутів та замикання множини ФЗ.

Побудовано математичне доведення коректності аксіоматики ФЗ табличних БД, а також її повноти через збіжність відношень синтаксичного та семантичного слідувань. В якості математичного апарата використані властивості теоретико-множинної конструкції обмеження функції за множиною.

Встановлено критерій повноти аксіоматики ФЗ Армстронга в термінах потужностей множини атрибутів та універсального домена.

Доведена незалежність складових аксіоматики Армстронга.

Побудована алгебра ФЗ, операції якої визначені відповідно до складових аксіоматики Армстронга. Розглянута підалгебра тривіальних ФЗ та множина, яка її породжує. Побудова цієї алгебри дозволяє формулювати результати щодо властивостей аксіоматики Армстронга на алгебраїчній мові.

Основні результати другого розділу наведені в табл. 2.2.

Таблиця 2.2

Основні результати розділу 2

№ з.п.	**Твердження**	**Стисле формулювання**	**Інтерпретація**
(1)	(2)	(3)	(4)
1.	Твердження 2.1.	$F \vdash X \to Y \Rightarrow F \vDash X \to Y$	Коректність аксіоматики Армстронга.
2.	Теорема 2.1.	$F \vDash X \to Y \Leftrightarrow F \vdash X \to Y$ (в припущенні, коли $\lvert D \rvert \geq 2$ і $\lvert R \rvert \geq 2$).	Повнота аксіоматики Армстронга.
3.	Теорема 2.2.	$F \vDash X \to Y \Leftrightarrow F \vdash X \to Y$ тоді і тільки тоді, коли в інтерпретації $\lvert D \rvert \geq 2$ або $\lvert R \rvert = 0$.	Критерій повноти аксіоматики Армстронга.

(1)	(2)	(3)	(4)
4.	Теорема 2.3.	Аксіоматика Армстронга є незалежною.	Видалення кожної складової зменшує породжуючу силу аксіоматики.
5.	Результат підрозділу 2.4	$A = \left\langle \left\{ X \to Y \right\}_{X, Y \in 2^R}, \{ R'^{(1)}_{\{A\}}, R''^{(2)}, \{A\} \to \varnothing^{(0)} \}_{A \in R} \right\rangle .$	Алгебра ФЗ.

РОЗДІЛ 3
АКСІОМАТИКА БАГАТОЗНАЧНИХ І ФУНКЦІОНАЛЬНИХ ЗАЛЕЖНОСТЕЙ

3.1 Аксіоматика багатозначних залежностей

Як і раніше, зафіксуємо множини R (скінченну множину атрибутів) і D (множину, з якої атрибути приймають значення в інтерпретаціях).

Наведемо означення БЗЗ в термінах роботи [76].

На таблиці t схеми R виконується БЗЗ $X \to\to Y$, якщо для двох довільних рядків s_1, s_2 таблиці t, які збігаються на множині атрибутів X, існує рядок $s_3 \in t$, який дорівнює об’єднанню обмежень рядків s_1, s_2 на множини атрибутів $X \bigcup Y$ і $R \setminus (X \bigcup Y)$ відповідно:

$$(X \to\to Y)(t) = true \overset{def}{\Leftrightarrow} \forall s_1, s_2 \in t(s_1 \mid X = s_2 \mid X \Rightarrow$$
$$\Rightarrow \exists s_3 \in t(s_3 = s_1 \mid (X \bigcup Y) \bigcup s_2 \mid R \setminus (X \bigcup Y))).$$

Отже, з семантичної точки зору БЗЗ – це параметричний предикат на таблицях схеми R, заданий двома (скінченними) множинами-параметрами атрибутів X, Y.

Структура таблиці t, на якій виконується БЗЗ $X \to\to Y$, може бути представлена за допомогою наступного відношення. Скажемо, що рядки s_1, s_2 таблиці t знаходяться у відношенні $=_X$, якщо вони збігаються на множині атрибутів X:

$$s_1 =_X s_2 \overset{def}{\Leftrightarrow} s_1 \mid X = s_2 \mid X .$$

Зрозуміло, що відношення $=_X$ є відношенням еквівалентності і тому розбиває множину рядків таблиці t на класи еквівалентності, які мають таке зображення:

$$[s]_{=_X} = \{s \mid X\} \otimes \pi_Y([s]_{=_X}) \otimes \pi_{R \setminus (X \bigcup Y)}([s]_{=_X}),$$

де s – довільний представник класу.

Скажемо, що таблиця $t(R)$ є моделлю множини БЗЗ G, якщо кожна БЗЗ $X \to\to Y \in G$ виконується на таблиці $t(R)$:

$$t(R) \textit{ модель } G \overset{def}{\Leftrightarrow} \forall (X \to\to Y)(X \to\to Y \in G \Rightarrow (X \to\to Y)(t) = true).$$

Скажемо, що БЗЗ $X \to\to Y$ семантично слідує з множини БЗЗ G за означенням, якщо на кожній таблиці $t(R)$, яка є моделлю множини БЗЗ G, виконується також БЗЗ $X \to\to Y$:

$$G \vDash X \to\to Y \overset{def}{\Leftrightarrow} \forall t(R)(t \textit{ модель } G \Rightarrow (X \to\to Y)(t) = true).$$

Лема 3.1 (*аксіома рефлексивності*): $\forall t(X \to\to Y)(t) = true$, де $Y \subseteq X$. □

Доведення. Розглянемо рядки s_1 і s_2 таблиці t, для яких виконується рівність $s_1 \mid X = s_2 \mid X$. Встановимо, що рядок $s_1 \mid (X \bigcup Y) \bigcup s_2 \mid R \setminus (X \bigcup Y)$ належить таблиці t, де R – як і раніше, схема таблиці t. Обмежимо обидві частини рівності $s_1 \mid X = s_2 \mid X$ за множиною Y: $(s_1 \mid X) \mid Y = (s_2 \mid X) \mid Y$. За властивістю оператора обмеження ($(U|Y)|Z = U|(Y \bigcap Z)$) згідно з [76, с. 24]) маємо $s_1|(X \bigcap Y) = s_2|(X \bigcap Y)$. Звідси і з умови $Y \subseteq X$ випливає $s_1|Y = s_2|Y$. З урахуванням дистрибутивності обмеження відносно об’єднання ($U \mid \bigcup_i X_i = \bigcup_i (U \mid X_i)$) (згідно з [76, с. 24]) маємо рівності:

$s_1 \mid (X \bigcup Y) \bigcup s_2 \mid R \setminus (X \bigcup Y) = s_1 \mid X \bigcup s_1 \mid Y \bigcup s_2 \mid R \setminus (X \bigcup Y) = s_2 \mid X \bigcup$
$\bigcup s_2 \mid Y \bigcup s_2 \mid R \setminus (X \bigcup Y) = s_2 \mid (X \bigcup Y \bigcup R \setminus (X \bigcup Y)) = s_2 \mid R = s_2$. □

Наслідок 3.1. $\varnothing \vDash X \to\to Y$ для $Y \subseteq X$. □

Лема 3.2. $\forall t(X \to\to Y)(t) = true$, де $X \bigcup Y = R$. □

Доведення. Розглянемо рядки s_1 і s_2 таблиці t, для яких виконується рівність $s_1 \mid X = s_2 \mid X$. Встановимо, що рядок $s_1 \mid (X \bigcup Y) \bigcup s_2 \mid R \setminus (X \bigcup Y)$ належить таблиці t, де R – як і раніше схема таблиці t. Врахувавши, що

$X \bigcup Y = R$, маємо рівності: $s_1 \mid (X \bigcup Y) \bigcup s_2 \mid R \setminus (X \bigcup Y) = s_1 \mid R \bigcup s_2 \mid \varnothing = s_1 \mid R =$ $= s_1 \in t$. □

Наслідок 3.2. $\varnothing \vDash X \to\to Y$ для $X \bigcup Y = R$. □

БЗЗ вигляду $X \to\to Y$, де $Y \subseteq X$ або $X \bigcup Y = R$, називається тривіальною.

Таким чином, тривіальні БЗЗ виконуються на довільній таблиці; іншими словами, тривіальні БЗЗ семантично слідують з порожньої множини БЗЗ.

Лема 3.3 (правило *повноти*).

$(X \to\to Y)(t) = true \Rightarrow (X \to\to R \setminus (X \bigcup Y))(t) = true$. □

Доведення. Нехай рядки s_1 і s_2 таблиці t такі, що виконується рівність $s_1 \mid X = s_2 \mid X$ і нехай виконується БЗЗ $X \to\to Y$. Встановимо, що рядок $s = s_1 \mid (X \bigcup R \setminus (X \bigcup Y)) \bigcup s_2 \mid R \setminus (X \bigcup R \setminus (X \bigcup Y))$ належить таблиці t. З рівностей $s_1 \mid X = s_2 \mid X$ і $R \setminus (X \bigcup R \setminus (X \bigcup Y)) = R \setminus (X \bigcup R \setminus Y) = Y \setminus X$, маємо: $s = s_1 \mid X \bigcup s_1 \mid (R \setminus (X \bigcup Y)) \bigcup s_2 \mid (Y \setminus X) = s_2 \mid X \bigcup s_2 \mid (Y \setminus X) \bigcup s_1 \mid (R \setminus (X \bigcup Y)) =$ $= s_2 \mid (X \bigcup Y \setminus X) \bigcup s_1 \mid (R \setminus (X \bigcup Y)) = s_2 \mid (X \bigcup Y) \bigcup s_1 \mid (R \setminus (X \bigcup Y))$. В силу довільності вибору рядків s_1 і s_2, а також з того, що на таблиці t виконується БЗЗ $X \to\to Y$ випливає, що рядок $s \in t$. □

Зауважимо, що за умови $Y \bigcap X \neq \varnothing$ зображення класів еквівалентності, на які відношення $=_X$ розбиває множину рядків таблиці t, можна уточнити:

$$[s]_{=_X} = \{s \mid X\} \otimes \pi_{Y \setminus X}([s]_{=_X}) \otimes \pi_{R \setminus (X \bigcup Y)}([s]_{=_X}).$$

Тоді очевидно, що рядок $s = s_1 \mid X \bigcup s_1 \mid (R \setminus (X \bigcup Y)) \bigcup s_2 \mid (Y \setminus X)$ з доведення леми 3.3 належить до того ж класу $[s]_{=_X}$, що і рядки s_1 і s_2.

Наслідок 3.3. $G \vDash X \to\to Y \Rightarrow G \vDash X \to\to R \setminus (X \bigcup Y)$. □

Доведення. Нехай БЗЗ $X \to\to Y$ семантично слідує з множини БЗЗ G. Нехай таблиця t – довільна модель множини БЗЗ G. Тоді за умовою $(X \to\to Y)(t) = true$. За правилом повноти $(X \to\to R \setminus (X \bigcup Y))(t) = true$. Отже, на довільній моделі множини БЗЗ G виконується БЗЗ $X \to\to R \setminus (X \bigcup Y)$.

Значить, за означенням семантичного слідування маємо $G \vDash X \to\to R \setminus (X \cup Y)$. □

Лема 3.4 (правило *поповнення*).

$(X \to\to Y)(t) = true \wedge Z \subseteq W \Rightarrow (X \cup W \to\to Y \cup Z)(t) = true$. □

Доведення. Розглянемо рядки s_1 і s_2 таблиці t, для яких виконується рівність $s_1 \mid (X \cup W) = s_2 \mid (X \cup W)$, тобто, $s_1 \mid X = s_2 \mid X$ і $s_1 \mid W = s_2 \mid W$. Встановимо, що рядок $s = s_1 \mid ((X \cup W) \cup (Y \cup Z)) \cup s_2 \mid R \setminus ((X \cup W) \cup (Y \cup Z))$ належить таблиці t, де R – як і раніше схема таблиці t. Оскільки $Z \subseteq W$, за властивістю дистрибутивності обмеження відносно об'єднання (як вказувалось вище [76, с. 24]) маємо: $s = s_1 \mid (X \cup Y) \cup s_1 \mid W \cup s_2 \mid R \setminus (X \cup Y \cup W)$. Враховуючи, що:

– виконується БЗЗ $X \to\to Y$, отже, рядок $s_3 = s_1 \mid (X \cup Y) \cup s_2 \mid R \setminus (X \cup Y)$ належить t; звідси маємо рівність $s_1 \mid (X \cup Y) = s_3 \mid (X \cup Y)$;

– з рівності $s_1 \mid W = s_2 \mid W$ та з побудови рядка s_3 випливає рівність $s_1 \mid W = s_3 \mid W$;

– з рівності $s_3 \mid R \setminus (X \cup Y) = s_2 \mid R \setminus (X \cup Y)$ (за побудовою рядка s_3) і включення $R \setminus (X \cup Y \cup W) \subseteq R \setminus (X \cup Y)$ випливає рівність $s_2 \mid R \setminus (X \cup Y \cup W) = s_3 \mid R \setminus (X \cup Y \cup W)$,

маємо: $s = s_3 \mid (X \cup Y) \cup s_3 \mid W \cup s_3 \mid R \setminus (X \cup Y \cup W) = s_3 \in t$. Отже, рядок s належить таблиці t. □

Наслідок 3.4. $G \vDash X \to\to Y \wedge Z \subseteq W \Rightarrow G \vDash X \cup W \to\to Y \cup Z$. □

Доведення. Нехай БЗЗ $X \to\to Y$ семантично слідує з множини БЗЗ G. Нехай далі таблиця t – довільна модель множини БЗЗ G. Тоді за умовою $(X \to\to Y)(t) = true$. За лемою 3.4 (правило поповнення) для $Z \subseteq W$ випливає $(X \cup W \to\to Y \cup Z)(t) = true$. Отже, на довільній моделі множини БЗЗ G виконується БЗЗ $X \cup W \to\to Y \cup Z$. Значить, за означенням семантичного слідування $G \vDash X \cup W \to\to Y \cup Z$. □

Лема 3.5 (правило *транзитивності*).

$$(X \to\to Y)(t) = true \wedge (Y \to\to Z)(t) = true \Rightarrow (X \to\to Z \setminus Y)(t) = true. \ \square$$

Доведення. Розглянемо рядки s_1 і s_2 таблиці t, для яких виконується рівність $s_1 | X = s_2 | X$. Тоді з умови $(X \to\to Y)(t) = true$ випливає, що рядок $s_3 = s_1 | (X \bigcup Y) \bigcup s_2 | R \setminus (X \bigcup Y)$ належить t. Оскільки для рядків s_1 і s_3 виконується рівність $s_1 | Y = s_3 | Y$ (за побудовою рядка s_3) то з виконання БЗЗ $Y \to\to Z$ випливає існування рядка $s_4 = s_1 | (Y \bigcup Z) \bigcup s_3 | R \setminus (Y \bigcup Z)$, який належить таблиці t. З рівності $Y \bigcup Z = Y \bigcup Z \setminus Y$, властивості дистрибутивності обмеження відносно об'єднання (як вказувалось вище [76, с. 24]) і рівності $s_1 | Y = s_3 | Y$ маємо: $s_4 = s_1 | (Y \bigcup Z \setminus Y) \bigcup s_3 | R \setminus (Y \bigcup Z) = s_1 | Y \bigcup s_1 | Z \setminus Y \bigcup$ $\bigcup s_3 | (R \setminus (Y \bigcup Z)) = s_3 | Y \bigcup s_1 | (Z \setminus Y) \bigcup s_3 | (R \setminus (Y \bigcup Z)) = s_1 | (Z \setminus Y) \bigcup s_3 | (Y \bigcup$ $\bigcup R \setminus (Y \bigcup Z)) = s_1 | (Z \setminus Y) \bigcup s_3 | (R \setminus (Z \setminus Y))$ (здійснено ланцюжок перетворень: $Y \bigcup R \setminus (Y \bigcup Z) = Y \bigcup R \setminus Z = R \setminus (Z \setminus Y)$).

Щоб довести, що БЗЗ $X \to\to Z \setminus Y$ виконується на таблиці t, покажемо, що рядок $s = s_1 | (X \bigcup Z \setminus Y) \bigcup s_3 | R \setminus (X \bigcup Z \setminus Y)$ належить таблиці t. За властивістю дистрибутивності обмеження відносно об'єднання з урахуванням рівності $s_1 | X = s_3 | X$ (за побудовою рядка s_3) маємо: $s = s_1 | X \bigcup s_1 | (Z \setminus Y) \bigcup$ $\bigcup s_3 | (R \setminus (X \bigcup Z \setminus Y)) = s_3 | X \bigcup s_1 | (Z \setminus Y) \bigcup s_3 | (R \setminus (X \bigcup Z \setminus Y)) = s_1 | (Z \setminus Y) \bigcup$ $\bigcup s_3 | (X \bigcup R \setminus (X \bigcup Z \setminus Y)) = s_1 | (Z \setminus Y) \bigcup s_3 | (X \bigcup R \setminus (Z \setminus Y))$. Оскільки $Z \setminus Y \bigcup$ $\bigcup R \setminus (Z \setminus Y) = R \supseteq X$ і $s_1 | X = s_3 | X$, то можемо записати $s = s_1 | (Z \setminus Y) \bigcup$ $\bigcup s_3 | (R \setminus (Z \setminus Y)) = s_4 \in t$. $\square$

Зауважимо, що за умови $Z \bigcap Y = \varnothing$ із виконання БЗЗ $X \to\to Y$ і $Y \to\to Z$ випливає виконання БЗЗ $X \to\to Z$.

Наслідок 3.5. Виконуються властивості відношення семантичного слідування:

1. $G \vDash X \to\to Y \wedge G \vDash Y \to\to Z \Rightarrow G \vDash X \to\to Z \setminus Y$;

2. $G \vDash X \to\to Y \wedge G \vDash Y \to\to Z \wedge Z \bigcap Y = \varnothing \Rightarrow G \vDash X \to\to Z$. $\square$

Доведення. ■ Доведемо перше твердження. Припустимо, що БЗЗ $X \to\to Y$, $Y \to\to Z$ семантично слідують з множини БЗЗ G. Нехай далі таблиця t – довільна модель множини БЗЗ G. Тоді за умовою $(X \to\to Y)(t) = true$ і $(Y \to\to Z)(t) = true$. Звідси за правилом транзитивності (лема 3.5) маємо $(X \to\to Z \setminus Y)(t) = true$. Отже, БЗЗ $X \to\to Z \setminus Y$ виконується на довільній моделі множини БЗЗ G, тобто $G \models X \to\to Z \setminus Y$. ▪

Доведення другого твердження проводиться безпосередньо. ▪□

Скажемо, що БЗЗ $X \to\to Y$ синтаксично слідує з множини БЗЗ G відносно схеми R ($G \vdash_R X \to\to Y$), якщо існує скінченна послідовність БЗЗ $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$, така, що $\varphi_m = X \to\to Y$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить G, або отримана за яким-небудь правилом виведення (повноти, поповнення, транзитивності) для БЗЗ з попередніх у цій послідовності БЗЗ φ_j, φ_k, $j, k < i$[7].

Послідовність $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$ назвемо доведенням, наслідуючи традиції математичної логіки [71].

Нехай задана деяка множина БЗЗ G. Замиканням $[G]_R$ називається множина БЗЗ, які синтаксично слідують з G відносно схеми R:

$$[G]_R \overset{def}{=} \{X \to\to Y \mid G \vdash_R X \to\to Y\}.$$

Для спрощення позначень далі параметр R явно вказувати не будемо.

В наведеній нижче лемі викладено деякі природні властивості замикання.

Лема 3.6. Виконуються властивості:

1) $G \subseteq [G]$ (*зростання*);

2) $[[G]] = [G]$ (*ідемпотентність*);

3) $G \subseteq H \Rightarrow [G] \subseteq [H]$ (*монотонність*).□

---

[7] Зауважимо, що правила повноти та поповнення застосовуються до однієї БЗЗ, а правило транзитивності – до двох.

Доведення. ■ Доведемо перше твердження. Нехай БЗЗ $X \to\to Y \in G$, тоді $G \vdash X \to\to Y$, оскільки послідовність одиничної довжини $X \to\to Y$ і є доведенням БЗЗ $X \to\to Y$. ▪

■ Доведемо друге твердження. За властивістю 1) маємо: $[G] \subseteq [[G]]$. Доведемо обернене включення $[[G]] \subseteq [G]$. Нехай $X \to\to Y$ – довільна БЗЗ, така, що $X \to\to Y \in [[G]]$. Тоді існує доведення $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$, де $\varphi_m = X \to\to Y$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить $[G]$, або отримана за яким-небудь правилом виведення для БЗЗ з попередніх у цій послідовності БЗЗ φ_j, φ_k, $j, k < i$. Утворимо нову послідовність БЗЗ за такими правилами:

– якщо φ_i є аксіома рефлексивності, то запишемо її без змін;

– якщо $\varphi_i \in [G]$, то за означенням замикання ця БЗЗ має скінченне доведення $\psi_1, ..., \psi_{l-1}, \psi_l$ з G; замість БЗЗ φ_i вставимо це її доведення;

– якщо φ_i отримана за яким-небудь правилом виведення з попередніх у цій послідовності БЗЗ φ_j, φ_k, $j, k < i$, то запишемо її без змін.

Очевидно, побудована таким чином послідовність є доведенням БЗЗ $X \to\to Y$ з G, тобто $G \vdash X \to\to Y$, а отже $X \to\to Y \in [G]$. ▪

■ Доведемо третє твердження. Нехай $X \to\to Y$ – довільна БЗЗ, така, що $X \to\to Y \in [G]$. Тоді існує доведення $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$, де $\varphi_m = X \to\to Y$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить G, або отримана за яким-небудь правилом виведення з попередніх у цій послідовності БЗЗ φ_j, φ_k, $j, k < i$. Оскільки множина БЗЗ G є підмножиною множини БЗЗ H, то, очевидно, що дане доведення є доведенням для БЗЗ $X \to\to Y$, виходячи з множини БЗЗ H: $X \to\to Y \in [H]$. ▪□

Таким чином, за термінологією [77] оператор $G \mapsto [G]$ є оператором замикання.

Зауважимо, що властивості оператора $G \mapsto [G]$, вказані в лемі 3.6, виконуються в аксіоматичних системах, де вводяться поняття аксіом та правил виведення (див., наприклад, [81]).

З описаних вище аксіоми рефлексивності і правил виведення будуються доведення похідних правил виведення для БЗЗ.

Лема 3.7 (правило *псевдотранзитивності*).

$\{X \to\to Y, Y \bigcup W \to\to Z\} \vdash X \bigcup W \to\to Z \setminus (Y \bigcup W)$. □

Доведення. Побудуємо доведення для БЗЗ $X \bigcup W \to\to Z \setminus (Y \bigcup W)$, виходячи з БЗЗ $X \to\to Y$ та $Y \bigcup W \to\to Z$:

1. $X \to\to Y$ (вихідна БЗЗ);
2. $X \bigcup W \to\to Y \bigcup W$ (з 1 за правилом поповнення);
3. $Y \bigcup W \to\to Z$ (вихідна БЗЗ);
4. $X \bigcup W \to\to Z \setminus (Y \bigcup W)$ (з 2 і 3 за правилом транзитивності). □

Лема 3.8 (правила *різниці*):

I. $\{X \to\to Y\} \vdash X \to\to Y \setminus X$;

II. $\{X \to\to Y \setminus X\} \vdash X \to\to Y$;

III. $\{X \to\to Y\} \vdash X \to\to R \setminus Y$.

Доведення.

I. Побудуємо доведення для БЗЗ $X \to\to Y \setminus X$:

1. $X \to\to X$ (аксіома рефлексивності);
2. $X \to\to Y$ (вихідна БЗЗ);
3. $X \to\to Y \setminus X$ (з 1 і 2 за правилом транзитивності). ▪

II. Побудуємо доведення для БЗЗ $X \to\to Y$:

1. $X \to\to Y \setminus X$ (вихідна БЗЗ);
2. $X \to\to Y$ (з 1 за правилом поповнення в результаті спрощення БЗЗ $X \bigcup (Y \bigcap X) \to\to Y \setminus X \bigcup (Y \bigcap X)$). ▪

III. Побудуємо доведення для БЗЗ $X \to\to R \setminus Y$:

1. $X \to\to Y$ (вихідна БЗЗ);
2. $X \to\to R \setminus (X \bigcup Y)$ (з 1 за правилом повноти);

3. $X \to\to R \setminus Y \bigcup X$ (з 2 за правилом поповнення в результаті спрощення БЗЗ $X \bigcup (X \setminus Y) \to\to R \setminus (X \bigcup Y) \bigcup (X \setminus Y)$). ▪□

Лема 3.9 (правило *адитивності*).

$\{X \to\to Y_1, X \to\to Y_2\} \vdash X \to\to Y_1 \bigcup Y_2$.□

Доведення. Побудуємо доведення для БЗЗ $X \to\to Y_1 \bigcup Y_2$:

1. $X \to\to Y_1$ (вихідна БЗЗ);
2. $X \to\to X \bigcup Y_1$ (з 1 за правилом поповнення);
3. $X \to\to Y_2$ (вихідна БЗЗ);
4. $X \bigcup Y_1 \to\to Y_2 \bigcup Y_1$ (з 3 за правилом поповнення);
5. $X \bigcup Y_1 \to\to R \setminus (X \bigcup Y_2 \bigcup Y_1)$ (з 4 за правилом повноти);
6. $X \to\to R \setminus (X \bigcup Y_2 \bigcup Y_1)$ (з 2 і 5 за правилом транзитивності; врахували що $(R \setminus (X \bigcup Y_2 \bigcup Y_1)) \setminus (X \bigcup Y_1) = R \setminus (X \bigcup Y_2 \bigcup Y_1)$, оскільки $R \setminus (X \bigcup Y_2 \bigcup Y_1) \bigcap$ $R \setminus (X \bigcup Y_2 \bigcup Y_1) \bigcap (X \bigcup Y_1) = \varnothing$);
7. $X \to\to X \bigcup Y_2 \bigcup Y_1$ (з 6 за правилом III леми 3.8);
8. $X \to\to Y_1 \bigcup Y_2$ (з 7 за правилом I леми 3.8). □

Наслідок 3.6. Для $n = 2, 3, \ldots$ має місце твердження:

$$\{X \to\to Y_1, ..., X \to\to Y_n\} \vdash X \to\to Y_1 \bigcup ... \bigcup Y_n .□$$

Доведення проводиться індукцією по n, виходячи з правила адитивності.

Лема 3.10 (правила *декомпозиції*).

I. $\{X \to\to Y_1, X \to\to Y_2\} \vdash X \to\to Y_1 \bigcap Y_2$;

II. $\{X \to\to Y_1, X \to\to Y_2\} \vdash X \to\to Y_1 \setminus Y_2$. □

Доведення.

I. Побудуємо доведення для БЗЗ $X \to\to Y_1 \bigcap Y_2$:

1. $X \to\to Y_1$ (вихідна БЗЗ);
2. $X \to\to R \setminus Y_1$ (з 1 за правилом III леми 3.8);
3. $X \to\to Y_2$ (вихідна БЗЗ);
4. $X \to\to R \setminus Y_2$ (з 3 за правилом III леми 3.8);

5. $X \to\to R \setminus (Y_1 \bigcap Y_2)$ (з 2 і 4 за правилом адитивності в результаті спрощення БЗЗ $X \to\to R \setminus Y_1 \bigcup R \setminus Y_2$);

6. $X \to\to Y_1 \bigcap Y_2$ (з 5 за правилом III леми 3.8). ▪

II. Побудуємо доведення для БЗЗ $X \to\to Y_1 \setminus Y_2$:

1. $X \to\to Y_2$ (вихідна БЗЗ);

2. $X \to\to R \setminus Y_2$ (з 1 за правилом III леми 3.8);

3. $X \to\to Y_1$ (вихідна БЗЗ);

4. $X \to\to Y_1 \setminus Y_2$ (з 2 і 3 за правилом I леми 3.10 в результаті спрощення $X \to\to R \setminus Y_2 \bigcap Y_1$). ▪□

Наслідок 3.7. Для $n = 2, 3, \ldots$ має місце твердження:

$$\{X \to\to Y_1, \ldots, X \to\to Y_n\} \vdash X \to\to Y_1 \bigcap \ldots \bigcap Y_n . \ \square$$

Доведення проводиться індукцією по *n*, виходячи з правила декомпозиції. □

Правила псевдотранзитивності, адитивності та декомпозиції наведені у роботах [5, 73].

3.2 Аксіоматика багатозначних і функціональних залежностей

Нехай задані множини ФЗ F і БЗЗ G. Скажемо, що таблиця $t(R)$ є моделлю множини $F \bigcup G$, якщо кожна залежність $\varphi \in F \bigcup G$ виконується на таблиці t:

$$t(R) \ \textit{модель} \ F \bigcup G \overset{def}{\Leftrightarrow} \forall \varphi \ (\varphi \in F \bigcup G \Rightarrow \varphi(t) = true) .$$

Лема 3.11. Для ФЗ та БЗЗ виконуються спільні правила виведення[8].

1. Правило *розширення* ФЗ до БЗЗ:

$(X \to Y)(t) = true \Rightarrow (X \to\to Y)(t) = true$.

2. $(X \to\to Z)(t) = true \wedge (Y \to Z')(t) = true \wedge Z' \subseteq Z \wedge Y \bigcap Z = \varnothing \Rightarrow$

$\Rightarrow (X \to Z')(t) = true$.

[8] Спільні правила виведення для ФЗ і БЗЗ наведені у [5].

Доведення. Доведення правила розширення ФЗ до БЗЗ наведено, наприклад, у монографії [76, с. 73]).

Доведемо спільне правило для ФЗ і БЗЗ з пункту 2. Нехай s_1 і s_2 – рядки таблиці t такі, що $s_1 | X = s_2 | X$, і нехай на цій таблиці виконується БЗЗ $X \to\to Z$; звідси випливає існування в таблиці t рядка $s_3 = s_1 | X \bigcup s_1 | Z \bigcup s_2 | R \setminus (X \bigcup Z)$. Нехай для множини $Z' \subseteq Z$ на таблиці t виконується ФЗ $Y \to Z'$, де $Y \bigcap Z = \varnothing$.

Покажемо спочатку, що для рядків s_2 і s_3 виконується рівність $s_2 | Y = s_3 | Y$. З рівностей $s_2 | X = s_3 | X$ (нагадаємо, що $s_1 | X = s_2 | X$ за припущенням і $s_1 | X = s_3 | X$ за побудовою рядка s_3) та $s_2 | R \setminus (X \bigcup Z) = s_3 | R \setminus (X \bigcup Z)$ (за побудовою рядка s_3) маємо рівність $s_2 | (R \setminus (X \bigcup Z) \bigcup X) = s_3 | (R \setminus (X \bigcup Z) \bigcup X)$, тобто[9] $s_2 | (R \setminus Z \bigcup X) =$ $= s_3 | (R \setminus Z \bigcup X)$; звідси з урахуванням включення $R \setminus Z \subseteq R \setminus Z \bigcup X$ маємо рівність $s_2 | (R \setminus Z) = s_3 | (R \setminus Z)$. Оскільки за умовою $Y \bigcap Z = \varnothing$, то $Y \subseteq R \setminus Z$, а отже, $s_2 | Y = s_3 | Y$.

За умовою $(Y \to Z')(t) = true$, отже, $s_2 | Z' = s_3 | Z'$. Оскільки $s_1 | Z = s_3 | Z$ (за побудовою рядка s_3), то з включення $Z' \subseteq Z$ випливає рівність $s_1 | Z' = s_3 | Z'$. Таким чином для рядків s_1 і s_2, які збігаються на множині атрибутів X, маємо рівність $s_1 | Z' = s_2 | Z'$. Отже ФЗ $X \to Z'$ виконується на таблиці t. □

Скажемо, що ФЗ або БЗЗ φ семантично слідує з множини залежностей $F \bigcup G$ відносно схеми R, якщо на кожній таблиці $t(R)$, яка є моделлю множини залежностей $F \bigcup G$, виконується також залежність φ:

$$F \bigcup G \vDash \varphi \overset{def}{\Leftrightarrow} \forall t(R)(t(R) - \textit{модель } F \bigcup G \Rightarrow \varphi(t) = true).$$

[9] Треба врахувати ланцюжок загальнозначних теоретико-множинних рівностей $R \setminus (X \bigcup Z) \bigcup Z = (R \setminus X \bigcap R \setminus Z) \bigcup Z = (R \setminus X \bigcup Z) \bigcap (R \setminus Z \bigcup Z) = (R \setminus X \bigcup Z) \bigcap R = R \setminus X \bigcup Z$ (де операція різниці має більший пріоритет).

Лема 3.12. Нехай H_1 і H_2 – множини залежностей (ФЗ або БЗЗ), і нехай T_1 і T_2 відповідно множини всіх їх моделей. Тоді виконується імплікація $H_1 \subseteq H_2 \Rightarrow T_1 \supseteq T_2$. □

Доведення. Нехай таблиця $t \in T_2$, тобто t – модель H_2. Очевидно, що таблиця t – модель H_1 і тому $t \in T_1$. □

Наслідок 3.8. Виконуються властивості відношення семантичного слідування:

1) $F \vDash \varphi \Rightarrow F \bigcup G \vDash \varphi$;

2) $G \vDash \varphi \Rightarrow F \bigcup G \vDash \varphi$. □

Доведення. ■ Доведемо перше твердження. Нехай $F \vDash \varphi$, тоді φ виконується на кожній моделі множини ФЗ F. Оскільки $F \subseteq F \bigcup G$, то за лемою 3.12 кожна модель множини $F \bigcup G$ є також моделлю множини F, отже, $F \bigcup G \vDash \varphi$. ▪

Доведення імплікації $G \vDash \varphi \Rightarrow F \bigcup G \vDash \varphi$ проводиться аналогічно. ▪□

Зрозуміло, що аналоги леми 3.12 та її наслідка 3.8 виконуються в теорії моделей аксіоматичних систем [71, 81] (це проявляється, зокрема, в тому, що в доведеннях леми та її наслідку по суті не використовувалася специфіка залежностей та таблиць).

З спільних правил виведення для ФЗ і БЗЗ, викладених в лемі 3.11, маємо наступний наслідок.

Наслідок 3.9. Виконуються властивості відношення семантичного слідування:

1. $F \vDash X \to Y \Rightarrow F \vDash X \to\to Y$;

2. $G \vDash X \to\to Z \wedge F \vDash Y \to Z' \wedge Z' \subseteq Z \wedge Y \bigcap Z = \varnothing \Rightarrow F \bigcup G \vDash X \to Z'$.

Доведення. ■ Доведемо перше твердження. Нехай ФЗ $X \to Y$ семантично слідує з множини ФЗ F. Нехай далі таблиця t – довільна модель множини ФЗ F. Тоді за умовою $(X \to Y)(t) = true$. За лемою 3.11 випливає $(X \to\to Y)(t) = true$. Отже, на довільній моделі множини ФЗ F виконується БЗЗ

$X \to\to Y$. Значить, за означенням семантичного слідування $F \vDash X \to\to Y$. ▪

■ Доведемо друге твердження. Нехай БЗЗ $X \to\to Z$ семантично слідує з множини БЗЗ G, а ФЗ $Y \to Z'$ семантично слідує з множини ФЗ F, тоді за наслідком 3.8 ці залежності семантично слідують з об'єднання множин ФЗ і БЗЗ $F \bigcup G$. Нехай далі таблиця t – довільна модель множини $F \bigcup G$. Тоді за умовою $(X \to\to Z)(t) = true$ і $(Y \to Z')(t) = true$. За лемою 3.11 для $Z' \subseteq Z$, де $Y \bigcap Z = \varnothing$ випливає $(X \to Z')(t) = true$. Отже, на довільній моделі множини $F \bigcup G$ виконується ФЗ $X \to Z'$. Значить, за означенням семантичного слідування $F \bigcup G \vDash X \to Z'$. ▪□

Лема 3.13. Виконуються властивості відношення семантичного слідування:

1) $F \vDash X \to Y \Rightarrow F \bigcup G \vDash X \bigcup Z \to Y \bigcup Z$, де $Z \subseteq R$;
   $F \bigcup G \vDash X \to Y \Rightarrow F \bigcup G \vDash X \bigcup Z \to Y \bigcup Z$, де $Z \subseteq R$;
2) $F \vDash X \to Y \wedge F \vDash Y \to Z \Rightarrow F \bigcup G \vDash X \to Z$;
   $F \bigcup G \vDash X \to Y \wedge F \bigcup G \vDash Y \to Z \Rightarrow F \bigcup G \vDash X \to Z$;
3) $G \vDash X \to\to Y \Rightarrow F \bigcup G \vDash X \to\to R \setminus (X \bigcup Y)$;
   $F \bigcup G \vDash X \to\to Y \Rightarrow F \bigcup G \vDash X \to\to R \setminus (X \bigcup Y)$
4) $G \vDash X \to\to Y \wedge Z \subseteq W \Rightarrow F \bigcup G \vDash X \bigcup W \to\to Y \bigcup Z$;
   $F \bigcup G \vDash X \to\to Y \wedge Z \subseteq W \Rightarrow F \bigcup G \vDash X \bigcup W \to\to Y \bigcup Z$
5) $G \vDash X \to\to Y \wedge G \vDash Y \to\to Z \Rightarrow F \bigcup G \vDash X \to\to Z \setminus Y$;
   $F \bigcup G \vDash X \to\to Y \wedge F \bigcup G \vDash Y \to\to Z \Rightarrow F \bigcup G \vDash X \to\to Z \setminus Y$
6) $F \vDash X \to Y \Rightarrow F \bigcup G \vDash X \to\to Y$;
   $F \bigcup G \vDash X \to Y \Rightarrow F \bigcup G \vDash X \to\to Y$;
7) $F \bigcup G \vDash X \to\to Z \wedge F \bigcup G \vDash Y \to Z' \wedge Z' \subseteq Z \wedge Y \bigcap Z = \varnothing \Rightarrow$
   $\Rightarrow F \bigcup G \vDash X \to Z'$. □

Доведення. ■ Доведемо перше твердження. Нехай $F \vDash X \to Y$, тоді за наслідком 2.2 (з розділу 2) маємо $F \vDash X \bigcup Z \to Y \bigcup Z$, де $Z \subseteq R$, звідси, за наслідком 3.8 випливає $F \bigcup G \vDash X \bigcup Z \to Y \bigcup Z$, де $Z \subseteq R$. ▪

■ Доведення імплікації $F \bigcup G \vDash X \to Y \Rightarrow F \bigcup G \vDash X \bigcup Z \to Y \bigcup Z$ проводиться аналогічно до доведення наслідку 2.2 (розділ 2). ▪

Доведення тверджень 2-7 леми 3.13 проводиться аналогічно до доведення твердження 1 з урахуванням наслідку 3.8 та відповідно:

– наслідку 2.3 (розділ 2) для доведення твердження 2;

– наслідку 3.3 для доведення твердження 3;

– наслідку 3.4 для доведення твердження 4;

– наслідку 3.5 для доведення твердження 5;

– наслідку 3.9 для доведення тверджень 6, 7. ▪□

Скажемо, що ФЗ або БЗЗ φ синтаксично слідує з об'єднання множин ФЗ і БЗЗ $F \bigcup G$ відносно схеми R ($F \bigcup G \vdash_R \varphi$), якщо існує скінченна послідовність ФЗ або БЗЗ $\varphi_1, \varphi_2, \ldots, \varphi_{m-1}, \varphi_m$, така, що $\varphi_m = \varphi$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності (для ФЗ або БЗЗ), або належить $F \bigcup G$, або отримана за яким-небудь правилом виведення (повноти (для БЗЗ), поповнення (для ФЗ або БЗЗ), транзитивності (для ФЗ або БЗЗ), спільних правил для ФЗ і БЗЗ) з попередніх у цій послідовності ФЗ або БЗЗ φ_j, φ_k, $j, k < i$.

Як і раніше, послідовність $\varphi_1, \varphi_2, \ldots, \varphi_{m-1}, \varphi_m$ назвемо доведенням φ з об'єднання множин $F \bigcup G$.

Нехай задані деякі множини F і G ФЗ та БЗЗ відповідно. Замикання $[F \bigcup G]_R$ – це множина усіх ФЗ і БЗЗ, які синтаксично слідують з $F \bigcup G$ відносно схеми R:

$$[F \bigcup G]_R \overset{def}{=} \{\varphi \mid F \bigcup G \vdash \varphi\}.$$

Лема 3.14. Виконуються властивості:

1) $F \bigcup G \subseteq [F \bigcup G]$ (*зростання*);

2) $[[F \bigcup G]] = [F \bigcup G]$ (*ідемпотентність*);

3) $F' \bigcup G' \subseteq F \bigcup G \Rightarrow [F' \bigcup G'] \subseteq [F \bigcup G]$[10] (*монотонність*);

---

[10] Зауважимо, що оскільки множини ФЗ і БЗЗ не перетинаються, то включення $F' \bigcup G' \subseteq F \bigcup G$ еквівалентне кон'юнкції включень вигляду $F' \subseteq F \wedge G' \subseteq G$.

4) $[F]\subseteq[F\bigcup G]$, $[G]\subseteq[F\bigcup G]$;

5) $[F]\bigcup[G]\subseteq[F\bigcup G]$. □

Доведення. Доведення тверджень 1-3 проводиться аналогічно до доведення відповідних пунктів леми 3.6.

■ Доведемо твердження 4. Дійсно, за твердженням 3 леми 3.14 з включень $F\subseteq F\bigcup G$ і $G\subseteq F\bigcup G$ випливають відповідно включення $[F]\subseteq[F\bigcup G]$ і $[G]\subseteq[F\bigcup G]$. ▪

■ Твердження 5 випливає безпосередньо з твердження 4: $[F]\subseteq[F\bigcup G]\wedge$ $\wedge[G]\subseteq[F\bigcup G]\Rightarrow[F]\bigcup[G]\subseteq[F\bigcup G]$. ▪□

З тверджень 1-3 леми 3.14 випливає, що оператор $F\bigcup G\mapsto[F\bigcup G]_R$ є оператором замикання.

Вкажемо ще одне спільне правило для БЗЗ і ФЗ [5]:

$$\{X\to\to Y, X\bigcup Y\to Z\}\vdash X\to Z\setminus Y.$$

Замиканням $[X]_{F\bigcup G,R}$ множини атрибутів X (відносно множини залежностей $F\bigcup G$ і схеми R) називається сім'я усіх правих частин БЗЗ, які синтаксично слідують з множини $F\bigcup G$:

$$[X]_{F\bigcup G,R}\overset{def}{=}\{Y\mid X\to\to Y\in[F\bigcup G]_R\}.$$

Очевидно, що $[X]_{F\bigcup G,R}\neq\varnothing$, оскільки, наприклад, $X\in[X]_{F\bigcup G,R}$, (бо $X\to\to X$, $X\to X$ є аксіомами рефлексивності); можна посилити останнє твердження: насправді виконується включення $2^X\subseteq[X]_{F\bigcup G,R}$, де 2^X – булеан множини атрибутів X.

Нехай $[X]_F$ – замикання множини атрибутів X відносно множини ФЗ F (див. розділ 2). Зауважимо, що за означенням $[X]_F\subseteq R$.

Лема 3.15. Виконуються властивості:

1) $Y\subseteq[X]_F\Rightarrow Y\in[X]_{F\bigcup G,R}$;

2) $[X]_{F\bigcup G,R}=[[X]_F]_{F\bigcup G,R}$. □

Доведення. ■ Для доведення першого твердження леми побудуємо доведення БЗЗ $X \to\to Y$, виходячи з множини залежностей $F \bigcup G$. Дійсно, маємо:

1. Доведення ФЗ $X \to [X]_F$ з F (твердження 2 леми 2.9 з розділу 2);

2. $[X]_F \to Y$ (аксіома рефлексивності для ФЗ; нагадаємо, що за припущенням $Y \subseteq [X]_F$);

3. $X \to Y$ (з 1 і 2 за правилом транзитивності для ФЗ);

4. $X \to\to Y$ (з 3 за правилом розширення ФЗ до БЗЗ).

Отже, за означенням замикання $[X]_{F \bigcup G,R}$ множина Y йому належить. ▪

■ Доведемо друге твердження. Нехай $Y \in [X]_{F \bigcup G,R}$, покажемо, що $Y \in [[X]_F]_{F \bigcup G,R}$. За означенням замикання $[X]_{F \bigcup G,R}$ існує доведення БЗЗ $X \to\to Y$ з множини залежностей $F \bigcup G$. Побудуємо доведення БЗЗ $[X]_F \to\to Y$ з множини залежностей $F \bigcup G$.

1. $[X]_F \to\to X$ (аксіома рефлексивності для БЗЗ; нагадаймо, що $X \subseteq [X]_F$ згідно твердження 1 леми 2.9 з розділу 2);

2. Доведення БЗЗ $X \to\to Y$ з $F \bigcup G$, яке існує за припущенням;

3. $[X]_F \to\to Y \setminus X$ (з 1 і 2 за правилом транзитивності для БЗЗ);

4. $[X]_F \to\to Y$ (з 3 за правилом поповнення для БЗЗ в результаті спрощення БЗЗ $[X]_F \bigcup (X \bigcap Y) \to\to Y \setminus X \bigcup (X \bigcap Y)$; дійсно, $Y \setminus X \bigcup (X \bigcap Y) = Y$; $[X]_F \bigcup (X \bigcap Y) = [X]_F$, оскільки $X \bigcap Y \subseteq [X]_F$).

Отже, маємо належність $Y \in [[X]_F]_{F \bigcup G,R}$. ▪

Нехай тепер $Y \in [[X]_F]_{F \bigcup G,R}$, покажемо, що $Y \in [X]_{F \bigcup G,R}$. За означенням замикання $[[X]_F]_{F \bigcup G,R}$ існує доведення БЗЗ $[X]_F \to\to Y$ з множини залежностей $F \bigcup G$. Побудуємо доведення БЗЗ $X \to\to Y$ з множини залежностей $F \bigcup G$.

1. Доведення ФЗ $X \to [X]_F$ з F (твердження 2 леми 2.9 з розділу 2);

2. $X \to\to [X]_F$ (з 1 за правилом розширення ФЗ до БЗЗ);

3. Доведення БЗЗ $[X]_F \to\to Y$ з множини $F \bigcup G$, яке існує за припущенням;

4. $X \to\to Y \setminus [X]_F$ (з 2 і останньої БЗЗ в послідовності доведення пункту 3 за правилом транзитивності для БЗЗ);

5. $[X]_F \to [X]_F \bigcap Y$ (аксіома рефлексивності для ФЗ);

6. $X \to [X]_F \bigcap Y$ (з останньої БЗЗ в послідовності доведення пункту 1 і з пункту 5 за правилом транзитивності для ФЗ);

7. $X \to\to [X]_F \bigcap Y$ (з 6 за правилом розширення ФЗ до БЗЗ);

8. $X \to\to Y$ (з 4 і 7 за правилом адитивності для БЗЗ в результаті спрощення правої частин БЗЗ $X \to\to (Y \setminus [X]_F) \bigcup ([X]_F \bigcap Y)$).

Отже, $Y \in [X]_{F \bigcup G,R}$. ▪□

Зауважимо, що оператор $X \mapsto [X]_{F \bigcup G,R}$ не є оператором замикання; для обґрунтування достатньо вказати, що даний оператор не володіє властивістю ідемпотентності (поняття множини атрибутів і замикання множини атрибутів відносно об’єднання множин ФЗ і БЗЗ $F \bigcup G$ та схеми R мають різну природу, тому вираз $[[X]_{F \bigcup G}]_{F \bigcup G,R}$ просто не має сенсу).

Базисом $[X]^{bas}_{F \bigcup G,R}$ множини атрибутів X відносно множини залежностей $F \bigcup G$ і схеми R називається підсім’я замикання $[X]_{F \bigcup G,R}$, така, що:

1) $\forall W (W \in [X]^{bas}_{F \bigcup G,R} \Rightarrow W \neq \varnothing)$ (тобто базис містить тільки непорожні множини атрибутів);

2) $\forall W_i W_j (W_i W_j \in [X]^{bas}_{F \bigcup G,R} \wedge W_i \neq W_j \Rightarrow W_i \bigcap W_j = \varnothing)$ (тобто різні множини атрибутів базису не перетинаються);

3) $\forall Y (Y \in [X]_{F \bigcup G,R} \Rightarrow \exists \Im (\Im \subseteq [X]^{bas}_{F \bigcup G,R} \wedge \Im$ – скінченна $\wedge Y = \bigcup_{W \in \Im} W$) (тобто кожна множина атрибутів з замикання $[X]_{F \bigcup G,R}$ є скінченним об’єднанням деяких множин атрибутів з базису).

Лема 3.16. Виконуються властивості:

1) $\bigcup_{W \in [X]^{bas}_{F \bigcup G, R}} W = R$ для $X \subseteq R$ (тобто базис є розбиттям множини атрибутів R);

2) $A \in [X]_F \Rightarrow \{A\} \in [X]^{bas}_{F \bigcup G, R}$. □

Доведення. ■ Доведемо перше твердження. Покажемо спочатку, що $\bigcup_{Y \in [X]_{F \bigcup G, R}} Y = R$. Оскільки сім'я множин $[X]_{F \bigcup G, R}$ непорожня, то виберемо множину $Y \in [X]_{F \bigcup G, R}$, тобто $X \to\to Y \in [F \bigcup G]_R$, тоді за правилом повноти $X \to\to R \setminus (X \bigcup Y) \in [F \bigcup G]_R$; отже, $R \setminus (X \bigcup Y) \in [X]_{F \bigcup G, R}$. Врахувавши, що $X \in [X]_{F \bigcup G, R}$ ($X \to\to X$ є аксіомою рефлексивності), маємо $\bigcup_{Z \in [X]_{F \bigcup G, R}} Z \supseteq Y \bigcup X \bigcup R \setminus (X \bigcup Y) = R$. Залишається врахувати тривіальне включення $\bigcup_{Z \in [X]_{F \bigcup G, R}} Z \subseteq R$.

Отже, сім'я $[X]_{F \bigcup G, R}$ є покриттям схеми R. Для завершення доведення першого твердження леми 3.16 залишається скористатися пунктом 3 з означення базису. ▪

■ Доведемо друге твердження. Побудуємо доведення для БЗЗ $X \to\to \{A\}$, де $A \in [X]_F$:

1. Доведення ФЗ $X \to [X]_F$ з F (твердження 2 леми 2.9 з розділу 2);

2. $[X]_F \to \{A\}$ (за аксіомою рефлексивності для ФЗ; нагадаємо, що $A \in [X]_F$);

3. $X \to \{A\}$ (з 1 і 2 за правилом транзитивності для ФЗ);

4. $X \to\to \{A\}$ (з 3 за правилом розширення ФЗ до БЗЗ),

отже, $\{A\} \in [X]_{F \bigcup G, R}$.

Покажемо тепер, що $\{A\} \in [X]^{bas}_{F \bigcup G,R}$. Ідея доведення полягає в тому, що сінглітон $\{A\}$ представити у вигляді об'єднання непорожніх підмножин можна єдиним способом $\{A\} = \bigcup_{B=A}\{B\}$.

Більш формальне доведення проводиться від супротивного. Нехай $\{A\} \notin [X]^{bas}_{F \bigcup G,R}$. Згідно з пунктом 3 означення базису існують такі різні елементи базису $W_1,...,W_n$, що $\{A\} = W_1 \cup ... \cup W_n$. Оскільки всі $W_1,...,W_n$ непорожні і $W_1,...,W_n$ попарно не перетинаються, то, очевидно, що $n=1$ та $\{A\} = W_1$; прийшли до суперечності. ▪□

3.3 Коректність та повнота аксіоматики багатозначних і функціональних залежностей

Нехай φ – ФЗ або БЗЗ.

Твердження 3.1 (коректність аксіоматики БЗЗ і ФЗ). Якщо залежність φ синтаксично виводиться з множини залежностей $F \bigcup G$, то φ виводиться з $F \bigcup G$ семантично:

$$F \bigcup G \vdash \varphi \Rightarrow F \bigcup G \vDash \varphi. \ \square$$

Доведення. Доведення проводиться індукцією за довжиною доведення.

Базис індукції. Довжина доведення дорівнює одиниці. Тобто залежність φ є або тривіальною (ФЗ або БЗЗ) або $\varphi \in F \bigcup G$. В обох випадках (для першого треба скористатися наслідками 3.1-3.2 для БЗЗ або наслідком 2.1 з розділу 2 для ФЗ) виконується $F \bigcup G \vDash \varphi$. ▪

Індуктивний крок. Нехай $\varphi_1, \varphi_2,...,\varphi_{m-1}, \varphi_m$, $m \geq 2$, доведення залежності φ (ФЗ або БЗЗ), виходячи з множини $F \bigcup G$. Розглянемо всі можливі випадки для останнього елемента доведення φ_m, де φ_m – ФЗ або БЗЗ.

Випадок, коли φ_m є або тривіальною (ФЗ або БЗЗ) або належить $F \bigcup G$, розглядається повністю аналогічно як в базисі індукції. ▪

Нехай φ_m – ФЗ, яка отримується з ФЗ φ_i, де $i < m$ за правилом поповнення. Очевидно, що $F \bigcup G \vdash \varphi_i$, за індуктивним припущенням маємо $F \bigcup G \vDash \varphi_i$. Залишається скористатися пунктом 1 леми 3.13. ▪

Аналогічно розглядаються випадки, коли:

– φ_m – ФЗ, яка отримується з попередніх ФЗ доведення за правилом транзитивності (використовується пункт 2 леми 3.13);

– φ_m – БЗЗ, яка отримується з БЗЗ φ_i, де $i < m$, за правилами повноти або поповнення (використовуються пункт 3 леми 3.13 для правила повноти або пункт 4 леми 3.13 для правила поповнення);

– φ_m – БЗЗ, яка отримується з попередніх БЗЗ доведення за правилом транзитивності (використовується пункт 5 леми 3.13);

– φ_m – БЗЗ, яка отримується з ФЗ φ_i, де $i < m$, за правилом розширення ФЗ до БЗЗ (використовується пункт 6 леми 3.13);

– φ_m – ФЗ, яка отримується з попередніх БЗЗ φ_i і ФЗ φ_j за спільним правилом для ФЗ і БЗЗ (використовується пункт 7 леми 3.13). ▪□

Твердження 3.2 (повнота аксіоматики БЗЗ і ФЗ). Якщо залежність φ семантично виводиться з множини залежностей $F \bigcup G$, то φ виводиться з $F \bigcup G$ синтаксично за умови $|R| \geq 2$ та $|D| \geq 2$[11] (при розгляді моделей):

$$F \bigcup G \vDash \varphi \Rightarrow F \bigcup G \vdash \varphi. \ \square$$

Доведення. Звернемось до ідеї доведення, викладеної в роботі [5], яку ми реконструюємо і доповнимо. Доведення проводиться від супротивного. Нехай множина $F \bigcup G$ і залежність φ (ФЗ або БЗЗ) такі, що $F \bigcup G \vDash \varphi$, але не виконується $F \bigcup G \vdash \varphi$, тобто, $\varphi \notin [F \bigcup G]$.

[11] Подробиці див. далі.

Прийдемо до протиріччя (з $F \bigcup G \vDash \varphi$), побудувавши таку модель множини $F \bigcup G$, що на ній залежність φ (ФЗ або БЗЗ) не буде виконуватися.

Зафіксуємо в універсальному домені два різних елементи a та b. Нехай множина X є лівою частиною залежності φ (ФЗ або БЗЗ) і нехай замикання $[X]_F$ складається з атрибутів $A_1, A_2, ..., A_k$. Тоді за властивістю 2 леми 3.16 для $i = \overline{1,k}$ маємо $\{A_i\} \in [X]^{bas}_{F \cup G, R}$, тобто, базис $[X]^{bas}_{F \cup G, R}$ здійснює розбиття схеми R потужності n на множини $A_1 = W_1, A_2 = W_2, ..., A_k = W_k, W_{k+1}, ..., W_m$, де $m \leq n$. Задамо таблицю t наступним чином: кількість рядків дорівнює 2^{m-k}, на кожному з атрибутів $A \in [X]_F$ усі рядки приймають значення тільки з множини $\{a\}$, тобто, $s(A) = a$; на множинах W_i для $i = \overline{k+1, m}$ значення кожного рядка обираються або тільки з множини $\{a\}$ або тільки з множини $\{b\}$ (див. табл. 3.1).

Таблиця 3.1

Таблиця t з доведення твердження 2

$[X]_F$	W_{k+1}	W_{k+2}	...	W_{m-1}	W_m
a	a	a	...	a	a
a	a	a	...	a	b
...	...	...	...	...	...
a	b	b	...	b	b

■ Нехай φ – ФЗ виду $X \to Y$ така, що $X \to Y \notin [F \bigcup G]$. Покажемо, що ФЗ $X \to Y$ не виконується на таблиці t. Оскільки, за припущенням, $X \to Y \notin [F \bigcup G]$, врахувавши включення $[F] \subseteq [F \bigcup G]$ (властивість 4 леми 3.14), маємо $X \to Y \notin [F]$. Звідси випливає, що $Y \not\subseteq [X]_F \subset R$ (властивість 3 леми 2.9 з розділу 2), тобто, $Y \bigcap R \setminus [X]_F \neq \varnothing$. Тоді для довільного атрибута $A \in Y \bigcap R \setminus [X]_F$ за побудовою таблиці t знайдуться такі рядки s_1 і s_2, для яких $s_1(A) = a$ і $s_2(A) = b$, отже, $s_1 | Y \neq s_2 | Y$. Врахувавши рівність $s_1 | X = s_2 | X$ (за побудовою таблиці t), маємо $(X \to Y)(t) = false$. Отримали протиріччя з припущенням. ▪

■ Нехай φ – БЗЗ виду $X \rightarrow\rightarrow Y$ така, що $X \rightarrow\rightarrow Y \notin [F \bigcup G]$. Покажемо, що БЗЗ $X \rightarrow\rightarrow Y$ не виконується на таблиці t. Оскільки, за припущенням, $X \rightarrow\rightarrow Y \notin [F \bigcup G]$ то:

1) $Y \not\subseteq [X]_F$, оскільки тоді БЗЗ $X \rightarrow\rightarrow Y$ мала б доведення (а, отже, належала б замиканню $[F \bigcup G]$):

a. Доведення ФЗ $X \rightarrow [X]_F$ з F, а отже, з $F \bigcup G$ (пункт 2 леми 2.9 з розділу 2);

b. $[X]_F \rightarrow Y$ (аксіома рефлексивності для ФЗ; нагадаємо, що за припущенням $Y \subseteq [X]_F$);

c. $X \rightarrow Y$ (з пунктів a і b за правилом транзитивності для ФЗ);

d. $X \rightarrow\rightarrow Y$ (з пункту c за правилом розширення ФЗ до БЗЗ);

2) $Y \neq \bigcup W_i$ для деяких $1 \leq i \leq m$, оскільки $W_i \in [X]^{bas}_{F \bigcup G, R}$, тобто, БЗЗ $X \rightarrow\rightarrow \bigcup W_i \in [F \bigcup G]$;

3) зауважимо також, що $Y \neq \varnothing$, оскільки БЗЗ $X \rightarrow\rightarrow \varnothing$ є аксіомою рефлексивності і належить замиканню $[F \bigcup G]$;

4) залишається розглянути випадок, коли для деякого i, $k+1 \leq i \leq m$ виконується включення $Y \bigcap W_i \subset W_i$, де $Y \bigcap W_i \neq \varnothing$. Враховуючи, що для довільних рядків s_1 і s_2 виконується рівність $s_1 | X = s_2 | X$ (за побудовою таблиці t) виберемо s_1 і s_2 такі, що $range(s_1 | W_i) = \{a\}$ і $range(s_2 | W_i) = \{b\}$ [12]. За правилом декомпозиції (пункт 2 леми 3.10) маємо $\{X \rightarrow\rightarrow W_i, X \rightarrow\rightarrow Y\} \vdash X \rightarrow\rightarrow W_i \setminus Y$, звідси випливає існування рядка s_3, для якого $s_3 | W_i = s_1 | (W_i \setminus Y) \bigcup s_2 | (W_i \bigcap Y)$, тобто $s_3 | W_i$ містить значення з обох множин $\{a\}$ і $\{b\}$, що суперечить побудові таблиці t. Отже, БЗЗ $X \rightarrow\rightarrow Y$, яка не належить замиканню $[F \bigcup G]$, не виконується на таблиці t, отримали протиріччя з припущенням. ▪

■ Покажемо тепер, що таблиця t є моделлю множини $F \bigcup G$.

[12] $range(s_i | W_i)$ – множина значень обмеження рядка s_i за множиною атрибутів W_i.

Розглянемо спочатку довільну ФЗ $U \to Z \in F \subseteq F \bigcup G$, та покажемо, що $(U \to Z)(t) = true$, тобто для довільних рядків s_1 і s_2 імплікація $s_1 | U = s_2 | U \Rightarrow s_1 | Z = s_2 | Z$ виконується.

Для множини атрибутів U можливі два випадки.

По-перше, нехай $U \bigcap R \setminus [X]_F = \varnothing$, тобто, $U \subseteq [X]_F$; тоді для довільних рядків s_1 і s_2 за побудовою таблиці t виконується рівність $s_1 | U = s_2 | U$, тому треба перевірити рівність $s_1 | Z = s_2 | Z$. Цю рівність ми доведемо, довівши включення $Z \subseteq [X]_F$. Для цього розглянемо доведення ФЗ $X \to Z$, виходячи з множини ФЗ F:

1. Доведення ФЗ $X \to [X]_F$ з F (пункт 2 леми 2.9 з розділу 2);

2. $[X]_F \to U$ (аксіома рефлексивності; нагадаємо, що розглядається випадок $U \subseteq [X]_F$);

3. $X \to U$ (правило транзитивності застосовано до ФЗ $X \to [X]_F$, яка є останнім елементом доведення з пункту 1, та ФЗ $[X]_F \to U$ з пункту 2);

4. $U \to Z$ (елемент множини F);

5. $X \to Z$ (правило транзитивності застосовано до ФЗ з пунктів 3 та 4).

Отже, маємо $F \vdash X \to Z$, тобто $Z \subseteq [X]_F$, звідси випливає, що $s_1 | Z = s_2 | Z$. ▪

По-друге, нехай $U \bigcap R \setminus [X]_F \neq \varnothing$; тоді, якщо $Z \subseteq [X]_F$, то ФЗ $U \to Z$ виконується тривіально, що випливає з побудови таблиці t.

Нехай $Z \not\subseteq [X]_F$.

Спочатку покажемо, що ФЗ виду $U \to Z$, де $Z \bigcap W_i \neq \varnothing$ і при цьому $U \bigcap W_i = \varnothing$ для $k+1 \leq i \leq m$, не належить множині F.

Від супротивного. Припустимо, що $U \to Z \in F$ і існує i, $k+1 \leq i \leq m$, що за умови $Z \bigcap W_i \neq \varnothing$ виконується рівність $U \bigcap W_i = \varnothing$. Прийдемо до протиріччя, побудувавши доведення ФЗ, яка не виконується на таблиці t:

1. $U \to Z$ (за припущенням є елементом множини F);

2. $Z \to Z \bigcap W_i$ (аксіома рефлексивності для ФЗ);

3. $U \to Z \bigcap W_i$ (з 1 і 2 за правилом транзитивності для ФЗ);

4. $Z \bigcap W_i \to W_i$ (за побудовою таблиці t; нагадаймо, що $\forall A', A'' \in W_i (s(A') = s(A''))$);

5. $U \to W_i$ (з 3 і 4 за правилом транзитивності для ФЗ);

6. $X \to\to W_i$ (за побудовою таблиці t; нагадаємо, що $W_i \in [X]^{bas}_{F \bigcup G,R}$);

7. $X \to W_i$ (з 6 і 5 за спільним правилом для ФЗ і БЗЗ; нагадаємо, що $U \bigcap W_i = \varnothing$).

Отримали доведення ФЗ $X \to W_i$ для деякого i, $k+1 \le i \le m$, яка не виконується на таблиці t, що випливає з її побудови. Отже, ФЗ $U \to Z$, де $Z \bigcap W_i \ne \varnothing$ і при цьому $U \bigcap W_i = \varnothing$ не належить множині F. ▪

Враховуючи, що $\bigcup_{W \in [X]^{bas}_{F \bigcup G,R}} W = R$ (пункт 1 леми 3.16) запишемо множину Z у вигляді $Z = \bigcup_{i=1}^{m} (Z \bigcap W_i)$. Зафіксуємо i і покажемо, що ФЗ $U \to Z \bigcap W_i$ виконується на таблиці t. Розглянемо можливі випадки для $Z \bigcap W_i$:

1) якщо $Z \bigcap W_i \subseteq [X]_F$, то ФЗ $U \to Z \bigcap W_i$ виконується за побудовою таблиці t;

2) якщо $Z \bigcap W_i \subseteq W_i$ для $k+1 \le i \le m$, то $U \bigcap W_i \ne \varnothing$ як було показано вище. Побудуємо доведення для ФЗ $U \to Z \bigcap W_i$:

a) $U \to U \bigcap W_i$ (аксіома рефлексивності для ФЗ);

b) $U \bigcap W_i \to Z \bigcap W_i$ (за побудовою таблиці t; нагадаємо, що $\forall A', A'' \in W_i (s(A') = s(A''))$ і виконуються включення $U \bigcap W_i \subseteq W_i$, $Z \bigcap W_i \subseteq W_i$);

c) $U \to Z \bigcap W_i$ (з a) і b) за правилом транзитивності для ФЗ).

Отже, ФЗ $U \to Z \bigcap W_i$ для $1 \le i \le m$ виконується на таблиці t, звідси за наслідком 2.5 (розділ 2) до правила композиції для ФЗ виконується ФЗ $U \to \bigcup_{i=1}^{m} (Z \bigcap W_i)$, тобто ФЗ $U \to Z$. ▪

■ Розглянемо тепер довільну БЗЗ $U \to\to Z \in G \subseteq F \bigcup G$ та покажемо, що $(U \to\to Z)(t) = true$.

Спочатку покажемо, що БЗЗ виду $U \to\to Z$, де $Z \bigcap W_i \subset W_i$ $(Z \bigcap W_i \ne \varnothing)$ і при цьому $U \bigcap W_i = \varnothing$ для $k+1 \le i \le m$ не належить G.

Від супротивного. Припустимо, що $U \to\to Z \in G$ і існує i, $k+1 \le i \le m$, що за умови $Z \bigcap W_i \subset W_i$ $(Z \bigcap W_i \ne \varnothing)$ виконується рівність $U \bigcap W_i = \varnothing$. Прийдемо до протиріччя, побудувавши доведення БЗЗ, яка не виконується на таблиці t:

1. $U \to\to Z$ (за припущенням є елементом множини G);

2. $R \setminus W_i \to\to Z$ (з 1 за правилом поповнення для БЗЗ з урахуванням включення $R \setminus W_i \supseteq Z \setminus W_i$ в результаті спрощення ФЗ $U \bigcup R \setminus W_i \to\to Z \bigcup Z \setminus W_i$; нагадаємо, що розглядається випадок $U \bigcap W_i = \varnothing$);

3. $X \to\to W_i$ (за побудовою таблиці t; нагадаємо, що $W_i \in [X]^{bas}_{F \bigcup G,R}$);

4. $X \to\to R \setminus W_i$ (за правилом різниці з пункту 3 леми 3.8);

5. $X \to\to Z \bigcap W_i$ (з 4 і 2 за правилом транзитивності для БЗЗ в результаті спрощення БЗЗ $X \to\to Z \setminus (R \setminus W_i)$).

Враховуючи, що $Z \bigcap W_i \subset W_i$ (нагадаємо, що $Z \bigcap W_i \ne \varnothing$) прийшли до суперечності з припущенням, що W_i належить базису замикання $[X]^{bas}_{F \bigcup G,R}$. Отже, БЗЗ виду $U \to\to Z$, де $Z \bigcap W_i \subset W_i$ $(Z \bigcap W_i \ne \varnothing)$ і при цьому $U \bigcap W_i = \varnothing$ для $k+1 \le i \le m$, не належить множині G. ▪

З урахуванням властивості $\bigcup_{W \in [X]^{bas}_{F \bigcup G,R}} W = R$ (пункт 1 леми 3.16) запишемо множину Z у вигляді $Z = \bigcup_{i=1}^{m} (Z \bigcap W_i)$. Зафіксуємо i і покажемо, що

БЗЗ $U \to\to Z \cap W_i$ виконується на таблиці t. Розглянемо можливі випадки для $Z \cap W_i$:

1. $Z \cap W_i = \varnothing$, тоді $U \to\to \varnothing$ є аксіомою рефлексивності і виконується тривіально;

2. $Z \cap W_i = W_i$; тоді $U \to\to W_i$ виконується за побудовою таблиці t, так як таблиця t містить усі можливі значення на множинах атрибутів W_i і $R \setminus W_i$, $k+1 \le i \le m$);

3. $Z \cap W_i \subseteq [X]_F$; тоді $U \to\to Z \cap W_i$ виконується за побудовою таблиці t;

4. $Z \cap W_i \subset W_i$ для $k+1 \le i \le m$ і $Z \cap W_i \ne \varnothing$, тоді $U \cap W_i \ne \varnothing$, як було показано вище.

Покажемо, що $U \to\to Z \cap W_i$ виконується для даного випадку, тобто, для довільних рядків s_1 і s_2 таких, що $s_1 | U = s_2 | U$ існує рядок s_3, такий, що $s_3 = s_1 | U \cup s_1 | (Z \cap W_i) \cup s_2 | R \setminus (U \cup (Z \cap W_i))$. За побудовою таблиці t з умов $s_1 | U = s_2 | U$ і $U \cap W_i \ne \varnothing$ випливає рівність $s_1 | W_i = s_2 | W_i$. Обмежимо обидві частини цієї рівності за множиною Z: $(s_1 | W_i) | Z = (s_2 | W_i) | Z$. За властивістю оператора обмеження $(U|Y)|Z = U|(Y \cap Z)$ згідно з [76, с. 24]) маємо $s_1 | (Z \cap W_i) = s_2 | (Z \cap W_i)$. Таким чином, $s_3 = s_2 | U \cup s_2 | (Z \cap W_i) \cup$ $\cup s_2 | R \setminus (U \cup (Z \cap W_i)) = s_2 \in t$. Отже, БЗЗ $U \to\to Z \cap W_i$ виконується на таблиці t.

Оскільки, як було показано, $U \to\to W_i$ виконується для $1 \le i \le m$, за наслідком 3.6 леми 3.9 маємо $\{U \to\to Z \cap W_1, U \to\to Z \cap W_2, ...,$ $..., U \to\to Z \cap W_m\} \vdash U \to\to Z$. Отже, БЗЗ $U \to\to Z$ виконується на таблиці t. ▪□

Теорема 3.1. Відношення семантичного та синтаксичного слідувань для аксіоматики БЗЗ та ФЗ за умови $|D| \ge 2$ та $|R| \ge 2$ збігаються:

$$F \cup G \vDash \varphi \Leftrightarrow F \cup G \vdash \varphi. \ \square$$

Доведення випливає безпосередньо з тверджень 3.1 і 3.2.

Аналогічна теорема має місце і для аксіоматики БЗЗ.

3.4 Критерій повноти аксіоматики багатозначних залежностей

Аналіз доведення основного результату про збіжність відношень синтаксичного ($\vdash$) і семантичного ($\vDash$) слідувань для аксіоматики БЗЗ та ФЗ (теорема 3.1) показує, що воно проведено в припущенні: $|D|\geq 2$ та $|R|\geq 2$, тобто універсальний домен містить не менше 2 елементів та схема R має щонайменше 2 атрибути.

Перевіримо збіжність відношень $\vdash$ і $\vDash$ у випадках $|D|\leq 2$ або $|R|\leq 2$ і тим самим уточнимо умови, в межах яких аксіоматика багатозначних залежностей є повною.

Лема 3.17. Усі БЗЗ виконуються на порожній таблиці. БЗЗ вигляду $X \to\to \varnothing$, зокрема, БЗЗ вигляду $\varnothing \to\to \varnothing$ виконується на довільній таблиці. □

Доведення випливає безпосередньо з означень. □

Лема 3.18. Довільна таблиця є моделлю порожньої множини БЗЗ. □

Лема 3.19. Множина тривіальних БЗЗ замкнена відносно правил повноти, поповнення та транзитивності. □

Доведення проводиться безпосередньо перевіркою відповідних імплікацій.

■ Нехай $X \to\to Y$ – тривіальна БЗЗ, покажемо, що БЗЗ $X \to\to R \setminus (X \bigcup Y)$, яка отримана за правилом повноти, є тривіальною. Розглянемо випадки:

1. $Y \subseteq X$; тоді $R \setminus (X \bigcup Y) = R \setminus X$, маємо БЗЗ $X \to\to R \setminus X$, яка є тривіальною (нагадаємо, що тривіальними є БЗЗ вигляду $X \to\to Y$, де $Y \subseteq X$ або $X \bigcup Y = R$);
2. $X \bigcup Y = R$; тоді $R \setminus (X \bigcup Y) = R \setminus R = \varnothing$, маємо БЗЗ $X \to\to \varnothing$, яка є тривіальною.

Отже, імплікація $X \to\to Y$ – тривіальна$\Rightarrow X \to\to R \setminus (X \bigcup Y)$ – тривіальна, виконується. ▪

■ Нехай $Z \subseteq W$ і нехай $X \to\to Y$ – тривіальна БЗЗ. Покажемо, що БЗЗ $X \bigcup W \to\to Y \bigcup Z$, яка отримана за правилом поповнення, є тривіальною. Розглянемо випадки:

a) $Y \subseteq X$; очевидно, що $X \bigcup W \supseteq Y \bigcup Z$ і БЗЗ $X \bigcup W \to\to Y \bigcup Z$ є тривіальною;

b) $X \bigcup Y = R$, тоді $(X \bigcup W) \bigcup (Y \bigcup Z) = (X \bigcup Y) \bigcup W = R \bigcup W = R$, тобто, БЗЗ $X \bigcup W \to\to Y \bigcup Z$ є тривіальною.

Отже, імплікація $X \to\to Y$ – тривіальна $\Rightarrow$ $X \bigcup W \to\to Y \bigcup Z$ – тривіальна, виконується. ▪

■ Нехай $X \to\to Y$ і $Y \to\to Z$ – тривіальні БЗЗ. Покажемо, що БЗЗ $X \to\to Z \setminus Y$, яка отримана за правилом поповнення, є тривіальною. Розглянемо випадки:

a) $Y \subseteq X$ і $Z \subseteq Y$; враховуючи включення $Z \setminus Y = \varnothing \subseteq X$ маємо тривіальну БЗЗ $X \to\to \varnothing$;

b) $X \bigcup Y = R$ і $Z \subseteq Y$; випадок розглядається аналогічно до попереднього;

c) $Y \subseteq X$ і $Y \bigcup Z = R$; тоді $R = Y \bigcup Z \setminus Y \subseteq X \bigcup Z \setminus Y \subseteq R$, звідси маємо $X \bigcup Z \setminus Y = R$, отже, БЗЗ $X \to\to Z \setminus Y$ є тривіальною;

d) $X \bigcup Y = R$ і $Y \bigcup Z = R$; тоді відповідно маємо $R \setminus Y \subseteq X$ і $R \setminus Y = Z \setminus Y$, звідси випливає включення $Z \setminus Y \subseteq X$, тобто, БЗЗ $X \to\to Z \setminus Y$ є тривіальною.

Імплікація $X \to\to Y$, $Y \to\to Z$ – тривіальна $\Rightarrow$ $X \to\to Z \setminus Y$ – тривіальна, виконується. ▪□

Лема 3.20. Всі БЗЗ виконуються на довільній однорядковій таблиці. □

Доведення випливає безпосередньо з означень. □

Нехай G – множина БЗЗ, через $(G)_{ntr}$ (None TRivial) позначимо підмножину множини G, що отримується вилученням тривіальних БЗЗ.

Лема 3.21. Виконуються наступні твердження:

1) $G \vdash \varphi$, якщо $\varphi \in G$;

2) $G \vDash \varphi$, якщо $\varphi \in G$;

3) $G \vdash \varphi \Leftrightarrow (G)_{ntr} \vdash \varphi$;

4) $G \vDash \varphi \Leftrightarrow (G)_{ntr} \vDash \varphi$;

5) $G \vdash \varphi$, якщо φ – тривіальна БЗЗ;

6) $G \vDash \varphi$, якщо φ – тривіальна БЗЗ. □

Доведення випливає безпосередньо з означень відношень синтаксичного та семантичного слідувань. □

Залежність збіжності відношень синтаксичного та семантичного слідувань для аксіоматики БЗЗ при різних значеннях потужностей множин R та D вказана у табл. 3.2. Дана таблиця містить три рядки та три стовпчики. Комірки будемо позначати парами, перша компонента – номер рядка, друга – номер стовпчика. “+” (відповідно “–”) в комірці означає, що при вказаних припущеннях відношення $\vdash$ і $\vDash$ збігаються (не збігаються).

Таблиця 3.2

Всі варіанти потужностей множин R та D для аксіоматики БЗЗ

		1	2	3
		$\|R\|=0$	$\|R\|=1$	$\|R\|\geq 2$
1	$\|D\|=0$	+	+	–
2	$\|D\|=1$	+	+	–
3	$\|D\|\geq 2$	+	+	+

Твердження 3.3. Заповнення таблиці 3.2 для аксіоматики БЗЗ коректне. □

Доведення проводиться розглядом 9 випадків.

■ Випадок (1,1): $|D|=|R|=0$. Множина всіх БЗЗ в інтерпретації задається з урахуванням очевидної рівності $I_R=\{X \rightarrow\rightarrow Y \mid X,Y \in 2^R\}$. У випадку, що розглядається ($R=\varnothing$), $I_\varnothing=\{\varnothing \rightarrow\rightarrow \varnothing\}$, оскільки $2^\varnothing=\{\varnothing\}$. Крім того, множина БЗЗ при розгляді моделей може бути і порожньою.

$t_\varnothing$, t_ε – всі таблиці в інтерпретації; вони є моделями таких множин БЗЗ: $G=\varnothing$ (лема 3.18) або $G=\{\varnothing \rightarrow\rightarrow \varnothing\}$ (лема 3.17).

Покажемо, що відношення $\vdash$ і $\vDash$ співпадають; для цього розглянемо всі можливі випадки:

$$\varnothing \vdash \varnothing \twoheadrightarrow \varnothing \Leftrightarrow \varnothing \vDash \varnothing \twoheadrightarrow \varnothing;$$

$$\{\varnothing \twoheadrightarrow \varnothing\} \vdash \varnothing \twoheadrightarrow \varnothing \Leftrightarrow \{\varnothing \twoheadrightarrow \varnothing\} \vDash \varnothing \twoheadrightarrow \varnothing.$$

Ці еквівалентності випливають з того, що виконуються твердження: $\varnothing \vdash \varnothing \twoheadrightarrow \varnothing$ (пункт 5 леми 3.21), $\varnothing \vDash \varnothing \twoheadrightarrow \varnothing$ (пункт 6 леми 3.21), $\{\varnothing \twoheadrightarrow \varnothing\} \vdash \varnothing \twoheadrightarrow \varnothing$ (пункт 1 леми 3.21), $\{\varnothing \twoheadrightarrow \varnothing\} \vDash \varnothing \twoheadrightarrow \varnothing$ (пункт 2 леми 3.21). ▪

■ Випадок (2,1): $|D| = 1$, $R = \varnothing$. Тоді $t_{\varnothing}$, t_{ε} – усі можливі таблиці в інтерпретації. Розглядається аналогічно випадку (1,1). ▪

■ Випадок (3,1): $|D| \geq 2$, $R = \varnothing$. Розглядається аналогічно випадкам (1,1) і (2,1). ▪

■ Випадок (1,2): $|R| = 1$, $D = \varnothing$. $t_{\varnothing}$ – єдина можлива таблиця в інтерпретації. Покладемо для визначеності $R \overset{def}{=} \{A\}$, тоді множина можливих БЗЗ задається як: $I_R = \{\varnothing \twoheadrightarrow \varnothing, \varnothing \twoheadrightarrow \{A\}, \{A\} \twoheadrightarrow \varnothing, \{A\} \twoheadrightarrow \{A\}\}$. Покажемо, що для довільної множини БЗЗ G, яка належить булеану множини I_R, відношення $\vdash$ і $\vDash$ співпадають; для цього перевіримо еквівалентності:

$$G \vdash \varnothing \twoheadrightarrow \varnothing \Leftrightarrow G \vDash \varnothing \twoheadrightarrow \varnothing;$$

$$G \vdash \varnothing \twoheadrightarrow \{A\} \Leftrightarrow G \vDash \varnothing \twoheadrightarrow \{A\};$$

$$G \vdash \{A\} \twoheadrightarrow \varnothing \Leftrightarrow G \vDash \{A\} \twoheadrightarrow \varnothing;$$

$$G \vdash \{A\} \twoheadrightarrow \{A\} \Leftrightarrow G \vDash \{A\} \twoheadrightarrow \{A\}.$$

Оскільки БЗЗ $\varnothing \twoheadrightarrow \varnothing$, $\varnothing \twoheadrightarrow \{A\}$, $\{A\} \twoheadrightarrow \varnothing$, $\{A\} \twoheadrightarrow \{A\}$ є тривіальними, то ліві частини еквівалентностей випливають з пункту 5 леми 3.21, а праві – з пункту 6 леми 3.21. ▪

■ Випадок (2,2): $|D| = 1$, $|R| = 1$. Нехай для визначеності $R \overset{def}{=} \{A\}$ та $D \overset{def}{=} \{d\}$. В інтерпретації можливі таблиці: $t_{\varnothing}$ і $t_0 = \{\{\langle A, d\rangle\}\} = \begin{array}{|c|}\hline A \\ \hline\hline d \\ \hline\end{array}$. У цьому випадку відношення $\vdash$ і $\vDash$ також співпадають. Розглядається аналогічно випадку (2,1). ▪

■ Випадок (3,2): $|D| \geq 2$, $|R| = 1$. Розглядається аналогічно випадкам (1,2) і (2,2). ▪

■ Випадок (1,3): $D = \varnothing$, $|R| \geq 2$. $t_{\varnothing}$ – єдина можлива таблиця в інтерпретації. Покажемо, що відношення $\vdash$ і $\vDash$ не співпадають; для цього наведемо наступний контрприклад. Нехай для визначеності атрибут $A \in R$ і нехай $G = \varnothing$, покажемо, що еквівалентність $\varnothing \vdash \varnothing \rightarrow\rightarrow \{A\} \Leftrightarrow \varnothing \vDash \varnothing \rightarrow\rightarrow \{A\}$ не виконується.

Дійсно, відношення $\varnothing \vDash \varnothing \rightarrow\rightarrow \{A\}$ має місце згідно з лемою 3.17.

Покажемо, що відношення $\varnothing \vdash \varnothing \rightarrow\rightarrow \{A\}$ не має місця. Доведення проводиться від супротивного. Припустимо, існує доведення $\varphi_1, \varphi_2, ..., \varphi_m$ БЗЗ $\varnothing \rightarrow\rightarrow \{A\}$ з порожньої множини БЗЗ $G = \varnothing$, де кожна φ_i є або аксіомою, або отримана з попередніх БЗЗ за допомогою правил повноти, поповнення чи транзитивності, а $\varphi_m = \varnothing \rightarrow\rightarrow \{A\}$. Згідно з лемою 3.19 усі БЗЗ з доведення є тривіальними. Тоді БЗЗ $\varnothing \rightarrow\rightarrow \{A\}$ також є тривіальною, прийшли до протиріччя з умовою $|R| \geq 2$ ($\varnothing \bigcup \{A\} \neq R$). Отже, вказана еквівалентність не виконується. ▪

■ Випадок (2,3): $|D| = 1$, $|R| \geq 2$. Відношення $\vdash$ і $\vDash$ не співпадають. Контрприклад той самий, що і для випадку (1,3) (в доведенні істинності відношення $\varnothing \vDash \varnothing \rightarrow\rightarrow \{A\}$ замість леми 3.17 використовується лема 3.20). ▪

■ Випадок (3,3): $|D| \geq 2$, $|R| \geq 2$. Відношення $\vdash$ і $\vDash$ співпадають. Це випливає з доповненого, реконструйованого доведення результату про повноту аксіоматики БЗЗ (теорема 3.1). ▪□

Теорема 3.2. Відношення семантичного $\vDash$ та синтаксичного $\vdash$ слідувань для аксіоматики БЗЗ співпадають тоді і тільки тоді, коли $|R| \leq 1$ або $|R| \geq 2$ і при цьому $|D| \geq 2$. □

3.5 Критерій повноти аксіоматики багатозначних і функціональних залежностей

Лема 3.22. Виконується імплікація: $X \to Y$ – тривіальна $\Rightarrow X \to\to Y$ – тривіальна. □

Доведення випливає безпосередньо з означень тривіальних ФЗ і БЗЗ. □

Наслідок 3.10. Множина тривіальних ФЗ і БЗЗ замкнена відносно правил виведення аксіоматики ФЗ і БЗЗ. □

Доведення випливає безпосередньо з леми 2.12 (розділ 2) та леми 3.19 і леми 3.22.□

Зауважимо, що спільне для ФЗ і БЗЗ правило $\dfrac{X \to\to Z, Y \to Z'}{X \to Z'}$, де $Z' \subseteq Z,\ Y \cap Z = \varnothing$, не може бути застосовано до двох тривіальних залежностей; це забезпечується умовами $Y \cap Z = \varnothing$ і $Z' \subseteq Z$, з яких випливає, що $Y \cap Z' = \varnothing$, отже, ФЗ $Y \to Z'$ не є тривіальною.

Залежність збіжності відношень синтаксичного та семантичного слідування для аксіоматики БЗЗ та ФЗ при різних значеннях потужностей множин R та D вказана у табл. 3.3. Усі позначення ті ж самі, що і для табл. 3.2.

Таблиця 3.3

Всі варіанти потужностей множин R та D для аксіоматики БЗЗ та ФЗ

		1	2	3
		$\|R\|=0$	$\|R\|=1$	$\|R\|\geq 2$
1	$\|D\|=0$	+	–	–
2	$\|D\|=1$	+	–	–
3	$\|D\|\geq 2$	+	+	+

Твердження 3.4. Заповнення таблиці 3.3 для аксіоматики БЗЗ і ФЗ коректне. □

Доведення проводиться аналогічно до доведення твердження 3.3.

■ Випадок (1,1): $|D|=|R|=0$. Очевидно, що множина $F\bigcup G$ належить булеану множини $\{\varnothing\to\varnothing, \varnothing\to\to\varnothing\}$. $t_{\varnothing}$, t_{ε} – всі таблиці в інтерпретації; вони є моделями множини $F\bigcup G=\varnothing$ (згідно з лемою 3.18) та множини $F\bigcup G\in\{2^{\{\varnothing\to\varnothing,\ \varnothing\to\to\varnothing\}}\setminus\varnothing\}$, оскільки ФЗ $\varnothing\to\varnothing$ і БЗЗ $\varnothing\to\to\varnothing$ – тривіальні.

Покажемо, що відношення $\vdash$ і $\vDash$ співпадають, тобто, виконується еквівалентність $F\bigcup G\vdash\varphi\Leftrightarrow F\bigcup G\vDash\varphi$, де φ – ФЗ або БЗЗ.

a) Нехай $F\bigcup G=\varnothing$; доведення еквівалентностей $\varnothing\vdash\varnothing\to\varnothing\Leftrightarrow\varnothing\vDash\varnothing\to\varnothing$ і $\varnothing\vdash\varnothing\to\to\varnothing\Leftrightarrow\varnothing\vDash\varnothing\to\to\varnothing$ було розглянуто відповідно в доведеннях твердження 2.4 (розділ 2) і твердження 3.3. Отже, при $F\bigcup G=\varnothing$ еквівалентність $\varnothing\vdash\varphi\Leftrightarrow\varnothing\vDash\varphi$ виконується.

b) $F\bigcup G\in\{2^{\{\varnothing\to\varnothing,\ \varnothing\to\to\varnothing\}}\setminus\varnothing\}$; очевидно, що множина $(F\bigcup G)_{ntr}=\varnothing$. Згідно з пунктами 3 і 4 леми 3.21 і з доведення попереднього випадку а) маємо: $F\bigcup G\vdash\varphi\Leftrightarrow(F\bigcup G)_{ntr}\vdash\varphi\Leftrightarrow\varnothing\vdash\varphi\Leftrightarrow\varnothing\vDash\varphi\Leftrightarrow(F\bigcup G)_{ntr}\vDash\varphi\Leftrightarrow$ $\Leftrightarrow F\bigcup G\vDash\varphi$. Отже, $F\bigcup G\vdash\varphi\Leftrightarrow F\bigcup G\vDash\varphi$. ▪

■ Випадок (2,1): $|D|=1$, $R=\varnothing$. Тоді $t_{\varnothing}$, t_{ε} – усі можливі таблиці в інтерпретації. Розглядається аналогічно випадку (1,1). ▪

■ Випадок (3,1): $|D|\geq 2$, $R=\varnothing$. Розглядається аналогічно випадкам (1,1) і (2,1). ▪

■ Випадок (1,2): $|R|=1$, $D=\varnothing$. $t_{\varnothing}$ – єдина можлива таблиця в інтерпретації. Покажемо, що відношення $\vdash$ і $\vDash$ не співпадають; для цього наведемо наступний контрприклад. Нехай для визначеності $R\overset{def}{=}\{A\}$ і нехай $F\bigcup G=\varnothing$, покажемо, що еквівалентність $\varnothing\vdash\varnothing\to\{A\}\Leftrightarrow\varnothing\vDash\varnothing\to\{A\}$ не виконується.

Дійсно, відношення $\varnothing\vDash\varnothing\to\{A\}$ має місце згідно з лемою 2.10 (розділ 2).

Покажемо, що відношення $\varnothing \vdash \varnothing \to \{A\}$ не має місця. Доведення проводиться від супротивного. Припустимо, існує доведення $\varphi_1, \varphi_2, \ldots, \varphi_m$ ФЗ $\varnothing \to \{A\}$ з порожньої множини ФЗ і БЗЗ $F \bigcup G = \varnothing$, де кожна φ_i є або аксіомою, або отримана з попередніх ФЗ або БЗЗ за допомогою відповідних до типу залежностей правил виведення, а $\varphi_m = \varnothing \to \{A\}$. Згідно з наслідком 3.10 усі ФЗ і БЗЗ з доведення є тривіальними, але ФЗ $\varnothing \to \{A\}$ не є тривіальною, прийшли до протиріччя. ▪

■ Випадок (2,2): $|D|=1$, $|R|=1$. Нехай для визначеності $R \overset{def}{=} \{A\}$ та $D \overset{def}{=} \{d\}$. В інтерпретації можливі таблиці: $t_\varnothing$ і $t_0 = \{\{\langle A, d \rangle\}\} = \begin{array}{|c|}\hline A \\ \hline\hline d \\ \hline \end{array}$. У цьому випадку відношення $\vdash$ і $\vDash$ також не співпадають. Контрприклад той самий, що і в попередньому випадку (1,2). ▪

■ Випадок (3,2): $|D| \geq 2$, $|R|=1$. Розглядається аналогічно випадкам (1,2) і (2,2). ▪

■ Випадок (1,3): $D = \varnothing$, $|R| \geq 2$. $t_\varnothing$ – єдина таблиця в інтерпретації. Відношення $\vdash$ і $\vDash$ не співпадають; контрприклад той самий, що і для випадку (1,2). ▪

■ Випадок (2,3): $|D|=1$, $|R| \geq 2$. В інтерпретації розглядаються однорядкові таблиці або $t_\varnothing$. Відношення $\vdash$ і $\vDash$ не співпадають. Котрприклад той самий, що і для випадків (1,2), (2,2). ▪

■ Випадок (3,3): $|D| \geq 2$, $|R| \geq 2$. Відношення $\vdash$ і $\vDash$ співпадають. Саме про це говорить доповнене, реконструйоване доведення результату про повноту аксіоматики ФЗ і БЗЗ (теорема 3.1). ▪□

Наслідок 3.11. *Якщо аксіоматика ФЗ повна для заданих потужностей елементів пари* $\langle R, D \rangle$*, то аксіоматика БЗЗ для даної пари також є повною.* □

Доведення випливає тривіально з заповнення табл. 2.1 (розділ 2) і табл. 3.2.□

Теорема 3.3. Відношення семантичного $\vDash$ та синтаксичного $\vdash$ слідувань для аксіоматики БЗЗ і ФЗ співпадають тоді і тільки тоді, коли $|D| \geq 2$ або $|R| = 0$. □

Доведення випливає з заповнення табл. 3.3. □

3.6 Основні результати розділу

В третьому розділі розглядаються аксіоматика БЗЗ і аксіоматика ФЗ та БЗЗ в табличних базах даних. Для кожної з цих аксіоматик введено поняття семантичного і синтаксичного слідувань.

Розглянуті властивості замикання множини атрибутів відносно об'єднання множин ФЗ і БЗЗ, замикання об'єднання множин ФЗ і БЗЗ та базису для об'єднання цих множин.

Побудовано математичне доведення коректності аксіоматики ФЗ і БЗЗ.

Узагальнено і доповнено класичний результат щодо повноти аксіоматики ФЗ і БЗЗ в табличних базах даних.

Встановлено критерії повноти вказаних аксіоматик в термінах потужностей універсального домену D, який розглядається при інтерпретаціях, і схеми R, яка є параметром всіх побудов, бо розглядаються тільки таблиці, схеми яких є підсхемами вказаної схеми R.

Основні результати третього розділу представлені в табл. 3.4.

Таблиця 3.4

Основні результати розділу 3

№ п.п.	**Твердження**	**Стисле формулювання**	**Інтерпретація**
(1)	(2)	(3)	(4)
1.	Твердження 3.1.	$F \cup G \vdash \varphi \Rightarrow F \cup G \vDash \varphi$	Коректність аксіоматики БЗЗ і ФЗ.

(1)	(2)	(3)	(4)
2.	Теорема 3.1.	$F \cup G \vDash \varphi \Leftrightarrow F \cup G \vdash \varphi$ за умови $\mid D \mid \geq 2$ і $\mid R \mid \geq 2$.	Повнота аксіоматики БЗЗ і ФЗ.
3.	Теорема 3.2.	$G \vDash \varphi \Leftrightarrow G \vdash \varphi$ тоді і тільки тоді, коли $\mid R \mid \leq 1$ або $\mid R \mid \geq 2$ і при цьому $\mid D \mid \geq 2$.	Критерій повноти аксіоматики БЗЗ.
4.	Теорема 3.3.	$F \cup G \vDash \varphi \Leftrightarrow F \cup G \vdash \varphi$ тоді і тільки тоді, коли $\mid D \mid \geq 2$ або $\mid R \mid = 0$.	Критерій повноти аксіоматики БЗЗ і ФЗ.

РОЗДІЛ 4
ФРАГМЕНТ МАТЕМАТИЧНОЇ ТЕОРІЇ НОРМАЛІЗАЦІЇ: НОРМАЛЬНІ ФОРМИ 2-4 ПОРЯДКІВ

Різні підходи до понятійного апарату теорії нормалізації в реляційних БД призвели до різноманітності у визначеннях НФ в таких БД (див. розділ 1).

Побудуємо фрагмент математичної теорії нормалізації (стосовно 2-4 НФ для табличних БД), який би відповідав вимогам строгості для математичних теорій.

4.1 Суперключі, потенційні ключі, первинні та непервинні атрибути, повні та транзитивні ФЗ

Нехай F – множина ФЗ. Зафіксуємо схему R.

ФЗ $X \to Y$ поставимо у відповідність множину ФЗ:

$$\Phi_{X\to Y} = \{X' \to Y \mid X' \subset X \wedge Y \not\subseteq X'\}.$$

Тобто, множина $\Phi_{X\to Y}$ містить усі нетривіальні ФЗ, такі, що їх права частина співпадає з множиною Y, а ліва частина є власною підмножиною множини X.

Очевидно, що $\Phi_{X\to Y} = \{X' \to Y \mid X' \in 2^X \setminus \{X\}\}$ за умови $X \subseteq Y$; $\Phi_{X\to Y} = \{X' \to Y \mid X' \in \{U \bigcup V \mid U \in 2^{X\setminus Y} \setminus \{X\}, V \in 2^Y \setminus \{Y\}\}\}$, за умови $Y \subseteq X$ або множини атрибутів X, Y непорівнювані відносно теоретико-множинного включення $\subseteq$. Вище 2^Y – булеан множини Y.

Означення 4.1. ФЗ $X \to Y$ називається повною (відносно множини ФЗ F), якщо вона виконується на довільній таблиці t, яка є моделлю множини F, а жодна ФЗ з множини $\Phi_{X\to Y}$ не виконується на таблиці t:

$$X \to Y \text{ – } \textit{повна} \text{ (відносно множини ФЗ } F\text{)} \overset{def}{\Leftrightarrow} \forall t(t \text{ – модель}$$
$$F \Rightarrow (X \to Y)(t) = true \wedge \forall \psi(\psi \in \Phi_{X\to Y} \Rightarrow \psi(t) = false)). \;\square$$

Враховуючи, що відношення синтаксичного і семантичного слідувань для аксіоматики ФЗ збігаються за умови $|D|\geq 2$ (див. критерій повноти аксіоматики Армстронга, теорема 2.2 з розділу 2), для вказаного випадку можна використовувати синтаксичну форму означення 4.1:

$$X \to Y \text{ – } \textit{повна} \text{ (відносно множини ФЗ } F\text{)} \overset{def}{\Leftrightarrow}$$

$$\overset{def}{\Leftrightarrow} X \to Y \in [F] \wedge \forall \psi (\psi \in \Phi_{X \to Y} \Rightarrow \psi \notin [F]),$$

де $[F]$ – синтаксичне замикання множини ФЗ F.

Означення 4.2. Множина атрибутів $W \subseteq R$ називається суперключем (відносно множини ФЗ F), якщо ФЗ $W \to R$ виконується на довільній таблиці t, яка є моделлю множини F:

$$W \text{ – } \textit{суперключ} \text{ (відносно множини ФЗ } F\text{)} \overset{def}{\Leftrightarrow}$$

$$\overset{def}{\Leftrightarrow} \forall t (t \text{ – модель } F \Rightarrow (W \to R)(t) = true). \ \square$$

Таким чином, наприклад, вся множина атрибутів схеми R є суперключем, оскільки ФЗ $R \to R$ є тривіальною і, значить, виконується на довільній таблиці.

Як і вище для випадку $|D|\geq 2$ можна використовувати синтаксичну форму означення 4.2:

$$W \text{ – } \textit{суперключ} \text{ (відносно множини ФЗ } F\text{)} \overset{def}{\Leftrightarrow} W \to R \in [F].$$

Означення 4.3. Множина атрибутів $K \subseteq R$ називається потенційним ключем (відносно множини ФЗ F), якщо K є суперключем і ФЗ $K \to R$ є повною. □

Змістовно кажучи, потенційний ключ – це суперключ, ліву частину якого не можна зменшити, не втративши властивість бути суперключем.

Лема 4.1 (про існування потенційного ключа). Для довільної множини ФЗ F, заданої над схемою R, існує не менше одного потенційного ключа. □

Доведення. Дійсно, якщо тривіальна ФЗ $R \to R$ є повною, то множина атрибутів R є шуканим потенційним ключем; інакше існує така власна підмножина $X^1 \subset R$, що на довільній моделі множини ФЗ F виконується ФЗ

$X^1 \to R$. Якщо ФЗ $X^1 \to R$ є повною, то множина атрибутів X^1 є потенційним ключем, інакше існує така власна підмножина $X^2 \subset X^1 \subset R$, що $(X^2 \to R)(t) = true$, де t – довільна модель множини ФЗ F. Продовжуючи таким чином і враховуючи, що множина атрибутів R містить скінченну кількість атрибутів, на k-ому кроці (для певного $k \le |R|$) знайдеться така множина атрибутів X^k, що $(X^k \to R)(t) = true$ і ФЗ $X^k \to R$ є повною. За означенням 4.3 множина атрибутів X^k і є шуканим потенційним ключем. □

Задамо на множині $2^R \times 2^R$ (по суті, на множині всіх ФЗ, бо ФЗ є парою множин атрибутів), наступне бінарне відношення. Скажемо, що ФЗ $X' \to Y'$ і $X \to Y$ знаходяться у відношенні $\trianglelefteq$, якщо $X' \subseteq X$ та $Y' = Y$:

$$X' \to Y' \trianglelefteq X \to Y \overset{def}{\Leftrightarrow} X' \subseteq X \wedge Y' = Y.$$

Очевидно, що відношення $\trianglelefteq$ на множині $2^R \times 2^R$ (тобто на множині всіх ФЗ) задає частковий порядок. Дійсно, відношення $\trianglelefteq$ є:

1) рефлексивним, оскільки $X \to Y \trianglelefteq X \to Y \Leftrightarrow X \subseteq X$;

2) антисиметричним, оскільки з того, що $X' \to Y \trianglelefteq X \to Y$ і $X \to Y \trianglelefteq X' \to Y$ відповідно слідують включення $X' \subseteq X$ і $X \subseteq X'$, звідки $X' = X$;

3) транзитивним, оскільки з того, що $X'' \to Y \trianglelefteq X' \to Y$ і $X' \to Y \trianglelefteq X \to Y$ відповідно слідують включення $X'' \subseteq X'$ і $X' \subseteq X$, звідки $X'' \subseteq X$, а отже, $X'' \to Y \trianglelefteq X \to Y$.

Розглянемо частково впорядковану множину $\left\langle \Phi_{R \to R} \bigcup \{R \to R\} \bigcap [F], \trianglelefteq \right\rangle$, де, як і раніше, $\Phi_{R \to R} = \{X \to R \mid X \subset R\}$. Тоді її кожний мінімальний елемент вигляду $X \to R$ є повною ФЗ і задає потенційний ключ X. Наявність декількох мінімальних елементів визначає декілька потенційних ключів. По суті це ще одне визначення потенційного ключа.

Значення потенційних ключів при деяких значеннях потужностей множин R та D вказані у табл. 4.1. Дана таблиця містить три рядки та чотири стовпчики. Комірки позначаються парами, перша компонента – номер рядка, друга – номер стовпчика.

Таблиця 4.1

Потенційні ключі при певних значеннях потужностей множин R та D

		1	2	3	4
	$\lvert R\rvert$ / $\lvert D\rvert$	$\lvert R\rvert=0$	$\lvert R\rvert=1$ $R=\{A\}$	$\lvert R\rvert=2$ $R=\{A,B\}$	$\lvert R\rvert>2$
1	$\lvert D\rvert=0$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
2	$\lvert D\rvert=1$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
3	$\lvert D\rvert\geq 2$	$\varnothing$	Єдиний ключ: $\{A\}$	Єдиний ключ: $\{A\}$ або $\{B\}$ або $\{A,B\}$. Два ключі: $\{A\}$ і $\{B\}$	Ключі визначаються, виходячи з множини ФЗ F

Лема 4.2. Заповнення табл. 4.1 коректне. □

Доведення. Доведення проводиться розглядом кожного з випадків.

■ $|R|=0$ (випадки (1,1), (2,1), (3,1), тобто перший стовпчик); тоді, незалежно від потужності домена, $t_{\varnothing}$, t_{ε} – всі можливі таблиці в інтерпретації, бо атрибутів в даному випадку немає. Єдиною можливою ФЗ є $\varnothing\to\varnothing$, яка є тривіальною, звідси $\varnothing$ – потенційний ключ; ▪

■ $|R|\geq 1$, $|D|\leq 1$ (випадки (1,2), (1,3), (1,4), (2,2), (2,3), (2,4), тобто комірки перших двох рядків, окрім комірок першого стовпчика); в інтерпретаціях розглядаються не більше ніж однорядкові таблиці; оскільки ФЗ вигляду $\varnothing\to R$ виконується на порожній та однорядковій таблицях (згідно з лемами 2.10 і 2.13 з розділу 2), з очевидних включень $\varnothing\subset\{A\}\subseteq R$, де A – довільний (або єдиний) атрибут схеми R, маємо, що єдиним потенційним ключем є порожня множина $\varnothing$; ▪

■ $|R|=1$, $|D|\geq 2$ (випадок (3,2)); в інтерпретаціях не більше ніж одноатрибутні таблиці; нехай для визначеності $R=\{A\}$. Оскільки ФЗ $\varnothing\to\{A\}$ не виконується на таблиці з двох рядків, то таблиця має єдиний потенційний ключ $\{A\}$ (ФЗ $\{A\}\to\{A\}$ є тривіальною); ▪

■ $|R|=2$ і $|D|\geq 2$ (випадок (3,3)); нехай для визначеності $R=\{A,B\}$ і $d_1,d_2\in D$, де d_1,d_2 – різні. Оскільки ФЗ $\varnothing\to R$ не виконується на довільній дворядковій таблиці та множини $\{A\}$, $\{A,B\}$, $(\{B\},\{A,B\})$ не можуть бути одночасно потенційними ключами з огляду на мінімальність потенційних ключів, то множина потенційних ключів може мати один ключ: $\{A\}$ або $\{B\}$ або $\{A,B\}$, чи два: $\{A\}$ і $\{B\}$. Множина ключів залежить від множини ФЗ F; точніше кажучи: якщо $F=\varnothing$, то єдиний потенційний ключ $\{A,B\}$; якщо $\{A\}\to\{A,B\}\in F$ $(\{B\}\to\{A,B\}\in F)$, то єдиний потенційний ключ $\{A\}$ (відповідно $\{B\}$); якщо $\{\{A\}\to\{A,B\},\{B\}\to\{A,B\}\}\subseteq F$, то маємо два потенційних ключа $\{A\}$ і $\{B\}$; ▪

■ $|R|\geq 2$ і $|D|\geq 2$ (випадок (3,4)); множина потенційних ключів цілком визначається множиною ФЗ F, тобто в цьому випадку нічого не можна сказати про потенційні ключі. ▪□

Наслідок 4.1. Якщо $|R|=0$ або $|D|\leq 1$, то потенційним ключем буде тільки порожня множина $\varnothing$. □

Доведення випливає з заповнення табл. 4.1. □

Лема 4.3. Якщо $|R|>0$ і $|D|\geq 2$, то ФЗ виду $\{A\}\to Y$, де $A\in R$, $\varnothing\neq Y\subseteq R$ є повною за умови $\{A\}\to Y\in[F]$ (тобто виконання ФЗ $\{A\}\to Y$ на довільній моделі множини ФЗ F). □

Для доведення достатньо врахувати той факт, що ФЗ вигляду $\varnothing\to Y$, де $Y\subseteq R$ і $Y\neq\varnothing$ не виконується на довільній дворядковій таблиці. □

Означення 4.4. Атрибут $A\in R$ називається транзитивно залежним від множини атрибутів $X\subset R$ (відносно множини ФЗ F), якщо існує така множина атрибутів $Y\subset R$, що $A\notin X\bigcup Y$ і для довільної таблиці t, яка є моделлю множини ФЗ F, виконується $(X \to Y)(t) = true$, $(Y \to X)(t) = false$, $(Y \to \{A\})(t) = true$[13]:

$$\exists Y(Y \subset R \wedge A \in R \setminus (X \bigcup Y) \wedge \forall t(t - модель$$

$$F \Rightarrow (X \to Y)(t) = true \wedge (Y \to X)(t) = false \wedge (Y \to \{A\})(t) = true). \ \square$$

Для випадку $|D| \geq 2$ запишемо синтаксичну форму означення 4.4:

атрибут A *транзитивно залежний* від множини атрибутів $X \subset R$

(відносно множини ФЗ F) $\overset{def}{\Leftrightarrow} \exists Y(Y \subset R \wedge \ A \in R \setminus (X \bigcup Y) \wedge$

$\wedge X \to Y \in [F] \wedge Y \to X \notin [F] \wedge Y \to \{A\} \in [F])$.

Лема 4.4. Якщо K – потенційний ключ (відносно множини ФЗ F) і для деякого атрибута $A \in R$ ФЗ $K \to \{A\}$ не є повною, то атрибут A є транзитивно залежним від множини атрибутів K. □

Доведення. Розглянемо довільну таблицю t, яка є моделлю множини ФЗ F. Нехай ФЗ $K \to \{A\}$ не є повною, оскільки ФЗ $K \to \{A\}$, очевидно, виконується на таблиці t (з огляду на те, що K – ключ), тоді існує така власна підмножина $K' \subset K$, $\{A\} \not\subset K'$ (тобто $A \notin K'$), що виконуються ФЗ $K' \to \{A\}$ на таблиці t; але оскільки K – потенційний ключ, то ФЗ $K' \to K$ не виконується на таблиці t з огляду на мінімальність потенційного ключа K. Таким чином, з виконання ФЗ $K \to K'$, $K' \to \{A\}$ і невиконання ФЗ $K' \to K$ на таблиці t (з огляду на її довільність) за означенням 4.4 атрибут A є транзитивно залежним від потенційного ключа K. □

Означення 4.5. Атрибут $A \in R$ називається первинним (відносно множини ФЗ F), якщо A міститься в якому-небудь потенційному ключі[14]:

$$A - первинний \overset{def}{\Leftrightarrow} A \in \bigcup_{K \in \Delta} K,$$

де $\Delta = \Delta(F)$ – сім’я всіх потенційних ключів (відносно множини ФЗ F). □

[13] Означення наведено згідно з [73]. Дрібас у монографії [68] в означенні транзитивної ФЗ надає додаткову умову: $Y \not\subset X$. Тоді подальша лема 4.4 не є чинною, а в наведеному нижче означенні 3НФ потрібно вимагати, щоб таблиця задовольняла умовам 2НФ.

[14] Означення 5-6 наводяться відповідно до [67, 68, 73].

Означення 4.6. Атрибут $A \in R$ називається непервинним (відносно множини ФЗ F), якщо A не міститься в жодному з потенційних ключів:

$$A - \textit{непервинний} \overset{def}{\Leftrightarrow} A \in R \setminus \bigcup_{K \in \Delta} K,$$

де Δ – як і раніше, сім'я всіх потенційних ключів. □

4.2 Друга нормальна форма

Як і раніше фіксуємо схему R та множину ФЗ F.

Означення 4.7. Скажемо, що таблиця $t(R)$ знаходиться у другій НФ (2НФ), якщо, по-перше, таблиця $t(R)$ знаходиться у першій НФ (1НФ[15]), по-друге, таблиця t є моделлю множини ФЗ F та, по-третє, кожний непервинний атрибут характеризується повною ФЗ від кожного потенційного ключа, тобто:

$$\forall A \forall K (K \in \Delta \wedge A \in R \setminus \bigcup_{\widehat{K} \in \Delta} \widehat{K} \Rightarrow K \to \{A\} - \text{повна ФЗ}). \ \square \qquad (1)$$

З означення 4.7 випливає, що, якщо множина непервинних атрибутів порожня, то умова (1) цього означення виконується автоматично (відповідна імплікація тривіально істинна, оскільки умова $A \in R \setminus \bigcup_{\widehat{K} \in \Delta} \widehat{K}$ хибна).

Твердження 4.1 (достатні умови для знаходження у 2НФ для спеціальних випадків). Таблиця t знаходиться у 2НФ якщо $|R| \leq 2$ або $|D| \leq 1$. □

Доведення. Використовуючи результати з табл. 4.1 розглянемо усі можливі випадки для умов $|R| \leq 2$ (перші три стовпчики) або $|D| \leq 1$ (перші два рядки):

■ $|R| = 0$ або $|D| \leq 1$ (комірки (1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1)); тобто два перших рядки і комірка (3,1); згідно з табл. 4.1 єдиним потенційним ключем є $\varnothing$, тобто $\Delta = \{\varnothing\}$, звідки всі атрибути схеми R є непервинними; оскільки в цьому випадку розглядаються не більше ніж

[15] Оскільки означення 1НФ не підлягає повній формалізації, пропонуємо використовувати один з його варіантів згідно з [67], адаптований до термінології даної статті: «таблиця знаходиться у 1НФ тоді і тільки тоді, коли кожний її рядок містить тільки одне значення для кожного атрибуту».

однорядкові таблиці, то, очевидно, що довільна така таблиця буде моделлю множини ФЗ F; таким чином, залишається перевірити виконання умови (1), яка в даному випадку приймає вигляд:

$$\forall A(A \in R \Rightarrow \varnothing \to \{A\} \text{ – повна ФЗ}),$$

та, знову ж виконується очевидно; ▪

■ $|D| \geq 2$ і $|R| = 1$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 4.1), тоді множина $\{A\}$ є потенційним ключем (див. табл. 4.1), а оскільки схема взагалі не містить непервинних атрибутів, то таблиця, яка є моделлю множини ФЗ F, знаходиться у 2НФ; ▪

■ $|D| \geq 2$ і $|R| = 2$ (комірка (3,3)); нехай для визначеності $R = \{A, B\}$ (як в табл. 4.1). Нехай потенційним ключем є одноелементна множина, наприклад, $\{A\}$. Тоді ФЗ $\{A\} \to \{B\}$, де $\{B\}$ – єдиний непервинний атрибут, є повною згідно з лемою 4.3. Отже, таблиця t, яка є моделлю множини ФЗ F, знаходиться у 2НФ. ▪

Нехай тепер множиною потенційних ключів є або $\Delta = \{\{A\}, \{B\}\}$ або $\Delta = \{\{A, B\}\}$; тоді схема взагалі не містить непервинних атрибутів, а отже, умова 2НФ (1) для таблиці t, яка є моделлю множини ФЗ F, виконується автоматично. ▪□

Змістовна інтерпретація твердження 4.1: не більше ніж двоатрибутні таблиці або не більше ніж однорядкові таблиці знаходяться у 2НФ (за умови, що вони є моделями множини ФЗ F).

В наступній лемі Δ – сім'я всіх потенційних ключів, як і раніше.

Лема 4.5 (достатні умови для знаходження у 2НФ). Таблиця t, яка є моделлю множини ФЗ F, знаходиться у 2НФ, якщо виконується одна з наступних умов:

1) всі потенційні ключі одноелементні;

2) $\bigcup_{K\in\Delta} K = R$, тобто сім'я потенційних ключів є покриттям схеми R (еквівалентно: множина непервинних атрибутів порожня)[16]. □

Доведення. ■ Доведемо перше твердження. Оскільки при значеннях потужностей $|D|\leq 1$ або $|R|=0$ потужність потенційного ключа нульова (наслідок 4.1), нам потрібно розглянути випадок, коли $|D|\geq 2$ і $|R|>0$. Нехай всі потенційні ключі одноелементні. Згідно з лемою 4.3 усі непервинні атрибути повністю залежать від кожного потенційного ключа; отже, таблиця t знаходиться у 2НФ. ▪

■ Доведемо друге твердження. Оскільки за умови 2) в формулюванні схема не містить непервинних атрибутів, то умова 2НФ для таблиці t не порушується. ▪□

4.3 Третя нормальна форма

Як і раніше фіксуємо схему R та множину ФЗ F.

Задамо на множині 2^R наступне бінарне відношення. Скажемо, що множини атрибутів $X, Y \subseteq R$ знаходяться у відношенні $\doteqdot$ (відносно множини ФЗ F), якщо $F \models X \to Y$ і $F \models Y \to X$.

Лема 4.7. Відношення $\doteqdot$ є відношенням еквівалентності. □

Доведення. Дійсно, відношення $\doteqdot$ є:

1) рефлексивним, оскільки $X \doteqdot X \Leftrightarrow F \models X \to X$ (ФЗ $X \to X$ є тривіальною);

2) симетричним, оскільки з того, що $X \doteqdot Y$ випливає $Y \doteqdot X$;

3) транзитивним; покажемо, що виконується імплікація $X \doteqdot Y \wedge Y \doteqdot Z \Rightarrow X \doteqdot Z$. Розглянемо випадки:

---

[16] Формулювання умов наводиться згідно з [70].

а) $|D|\le 1$ (перші два рядки табл. 4.1); оскільки на порожній та однорядковій таблицях виконуються довільні ФЗ, то імплікація виконується очевидно;

б) $|R|=0$ (перший стовпчик табл. 4.1); єдиною можливою ФЗ є тривіальна ФЗ $\varnothing \to \varnothing$ отже, імплікація виконується тривіально;

в) $|D|\ge 2 \wedge |R|\ge 1$; враховуючи, що для даного випадку відношення синтаксичного слідування ($\vdash$) та семантичного слідування ($\models$) співпадають (див. теорему 2.2 про критерій повноти аксіоматики Армстронга), побудуємо доведення для ФЗ $X \to Z$:

1. $X \to Y$ (за припущенням $X \doteqdot Y$; отже, $F \models X \to Y$, звідси за критерієм повноти аксіоматики Армстронга маємо: $F \vdash X \to Y$);

2. $Y \to Z$ (за припущенням $Y \doteqdot Z$; отже, $F \models Y \to Z$, звідси за критерієм повноти аксіоматики Армстронга маємо: $F \vdash Y \to Z$);

3. $X \to Z$ (з 1 і 2 за правилом транзитивності).

Отже, $F \vdash X \to Z$, звідси випливає $F \models X \to Z$ (нагадаємо, що $|D|\ge 2$). Аналогічно будується доведення для ФЗ $Z \to X$, тобто, $F \vdash Z \to X$, отже $F \models Z \to X$. Звідси маємо: $X \doteqdot Z$. □

Таким чином, відношення $\doteqdot$ розбиває множину 2^R (булеан схеми R) на класи еквівалентності.

Наслідок 4.2. За умов $|D|\le 1$ або $|R|=0$ усі елементи множини $2^R = \{X \mid X \subseteq R\}$ належать до одного класу еквівалентності. □

Доведення випливає безпосередньо з доведення леми 4.6. □

Лема 4.7. Множина суперключів належить до одного класу еквівалентності. □

Доведення. Нехай множина атрибутів K є суперключем (відносно множини ФЗ F), тоді виконується ФЗ $K \to R$, тобто $F \models K \to R$. Враховуючи виконання тривіальної ФЗ $R \to K$ ($K \subseteq R$), тобто $F \models R \to K$, маємо $K \doteqdot R$. □

Позначимо клас еквівалентності суперключів $[R]_{\doteqdot}$, де схема R є представником класу.

Лема 4.8. Виконується імплікація:

$$\forall A, \forall X(\forall Y(A \in R \setminus (X \bigcup Y) \wedge (Y \to \{A\})(t) = true \Rightarrow X \doteqdot Y) \Rightarrow$$

$$\Rightarrow A \text{ не є транзитивно залежним від } X). \square$$

Доведення проводиться безпосередньо. □

Відомо, що наявність непервинних атрибутів, які є транзитивно залежними від потенційних ключів, призводить до виникнення аномалій (згідно визначення аномалій з [26]). З леми 4.8 випливає, що для вирішення вказаної проблеми достатньо вимагати, щоб ліві частини усіх ФЗ виду $X \to \{A\}$, де $\{A\}$ – непервинний атрибут, належали до класу $[R]_{\doteqdot}$ [17].

Означення 4.8. Скажемо, що таблиця $t(R)$ знаходиться у третій НФ (3НФ), якщо, по-перше, таблиця $t(R)$ знаходиться у 1НФ, по-друге, таблиця t є моделлю множини ФЗ F та, по-третє, для кожної нетривіальної ФЗ виду $X \to \{A\}$ такої, що $F \models X \to \{A\}$, де A – непервинний атрибут, множина атрибутів X є суперключем, тобто:

$$\forall X, \forall A(X \subset R \wedge A \in R \setminus (\bigcup_{K \in \Delta} K \cup X) \wedge$$

$$\wedge (X \to \{A\})(t) = true \Rightarrow (X \to R)(t) = true). \square \qquad (2)$$

З означення 4.8 випливає, що, якщо множина непервинних атрибутів порожня, то умова (2) цього означення виконується автоматично (відповідна імплікація тривіально істинна). Таким чином, доведена наступна лема.

Лема 4.9 (достатня умова для знаходження у 3НФ). Таблиця t, яка є моделлю множини ФЗ F, знаходиться у 3НФ якщо $\bigcup_{K \in \Delta} K = R$, тобто сім’я потенційних ключів є покриттям схеми R. □

Твердження 4.2 (достатні умови для знаходження у 3НФ для спеціальних випадків). Таблиця знаходиться у 3НФ якщо $|R| \leq 2$ або $|D| \leq 1$. □

Доведення. Враховуючи результати з табл. 4.1 розглянемо усі можливі випадки для умов $|R| \leq 2$ (перші три стовпчики) або $|D| \leq 1$ (перші два рядки):

---

[17] Саме такий підхід здійснюється у класичних алгоритмах, описаних у [68].

■ $|R|=0$ (комірки (1,1), (2,1), (3,1), тобто перший стовпчик); оскільки схема R не має непервинних атрибутів, то умови 3НФ виконуються автоматично (лема 4.9); ▪

■ $|R|\geq 1$ і $|D|\leq 1$ (комірки (1,2), (1,3), (1,4), (2,2), (2,3), (2,4); тобто перші два рядки); згідно з табл. 4.1 єдиним потенційним ключем є $\varnothing$, тобто $\Delta=\{\varnothing\}$, тоді будь-який елемент множини $2^R=\{X \mid X\subseteq R\}$ є суперключем; таким чином, умова (2) виконується для кожної відповідної ФЗ вигляду $X\to\{A\}$; оскільки в цьому випадку розглядаються не більше ніж однорядкові таблиці, то, очевидно, що довільна така таблиця буде моделлю множини ФЗ F і знаходиться у 3НФ; ▪

■ $|R|=1$, $|D|\geq 2$ (комірка (3,2)); нехай для визначеності $R=\{A\}$ (як в табл. 4.1), тоді множина $\{A\}$ є потенційним ключем (див. табл. 4.1), а оскільки схема не містить непервинних атрибутів, то за лемою 4.9 таблиця, яка є моделлю множини ФЗ F, знаходиться у 3НФ; ▪

■ $|R|=2$, $|D|\geq 2$ (комірка (3,3)); нехай для визначеності $R=\{A,B\}$ (як в табл. 4.1). Нехай потенційним ключем є одноелементна множина, наприклад, $\{A\}$, тоді B – єдиний непервинний атрибут. Запишемо усі елементи множини $2^R=\{X \mid X\subseteq R\}$: $\{\varnothing,\{A\},\{B\},\{A,B\}\}$. Оскільки умова (2) виключає з розгляду тривіальні ФЗ, то залишається перевірити лише імплікації:

1. $(\varnothing\to\{B\})(t)=true \Rightarrow (\varnothing\to\{A,B\})(t)=true$;
2. $(\{A\}\to\{B\})(t)=true \Rightarrow (\{A\}\to\{A,B\})(t)=true$.

В першій імплікації посилка є хибною, оскільки ФЗ $\varnothing\to\{B\}$ не виконуються на довільній дворядковій таблиці, отже, імплікація виконується.

Друга імплікація виконується, оскільки за припущенням множина $\{A\}$ є потенційним ключем. Отже, таблиця, яка є моделлю множини ФЗ F, знаходиться у 3НФ. ▪

Нехай тепер множиною потенційних ключів є або $\{\{A\},\{B\}\}$ або $\{(A,B\}\}$; тоді схема взагалі не містить непервинних атрибутів, а отже, за лемою 4.9 таблиця, яка є моделлю множини ФЗ F, знаходиться у 3НФ. ▪□

Змістовна інтерпретація твердження 4.2: не більше ніж двоатрибутні таблиці або не більше ніж однорядкові таблиці знаходяться у 3НФ (за умови, що вони є моделями множини ФЗ F).

Наслідок 4.3. $3НФ \Leftrightarrow 2НФ$ за умови $|R| \leq 2$ або $|D| \leq 1$.□

Доведення випливає з лем 4.1 і 4.2.□

Лема 4.10. Якщо таблиця t знаходиться у 3НФ, то вона знаходиться у 2НФ: $3НФ \Rightarrow 2НФ$.□

Доведення. З урахуванням наслідку 4.3 достатньо розглянути випадок, коли $|D| \geq 2$ і $|R| > 2$.

Нехай таблиця t, яка є моделлю множини ФЗ F, знаходиться у 3НФ. Доведення проводиться від супротивного. Припустимо таблиця t не знаходиться у 2НФ, тоді для деякого непервинного атрибута A знайдеться такий потенційний ключ K, що ФЗ $K \rightarrow \{A\}$ не є повною, тобто, існує така підмножина $K' \subset K$, що виконується ФЗ $K' \rightarrow \{A\}$. Це порушує виконання умови (2), оскільки в силу мінімальності потенційного ключа множина $K' \subset K$ не є суперключем. Прийшли до протиріччя. □

4.4 Нормальна форма Бойса-Кодда

Означення 4.9. Скажемо, що таблиця $t(R)$ знаходиться у НФ Бойса-Кодда (НФБК), якщо, по-перше, таблиця $t(R)$ знаходиться у 1НФ, по-друге, таблиця t є моделлю множини ФЗ F та, по-третє, для кожної нетривіальної ФЗ виду $X \rightarrow \{A\}$ такої, що $F \vDash X \rightarrow \{A\}$, множина атрибутів X є суперключем, тобто:

$$\forall X, \forall A\,(X \subset R \wedge A \in R \setminus X \wedge (X \rightarrow \{A\})(t) = true \Rightarrow (X \rightarrow R)(t) = true).\square \quad (3)$$

Твердження 4.3 (достатні умови для знаходження у НФБК для спеціальних випадків). Таблиця знаходиться у НФБК якщо $|R| \le 2$ або $|D| \le 1$. □

Доведення. Враховуючи результати з табл. 4.1 розглянемо усі можливі випадки для умов $|R| \le 2$ (перші три стовпчики) або $|D| \le 1$ (перші два рядки).

■ $|R| = 0$ або $|D| \le 1$ (комірки (1,1), (2,1), (3,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4); тобто перший стовпчик і перші два рядки); згідно з наслідком 4.2 усі елементи множини $2^R = \{X \mid X \subseteq R\}$ належать до одного класу еквівалентності – класу суперключів $[R]_{\equiv}$, де схема R є представником класу. Отже, умови НФБК виконуються автоматично і таблиця, яка є моделлю множини ФЗ F, знаходиться у НФБК; ▪

■ $|R| = 1$, $|D| \ge 2$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 4.1), тоді множина $\{A\}$ є потенційним ключем (див. табл. 4.1). Імплікація з умови (3) для єдиної нетривіальної ФЗ $\varnothing \to \{A\}$ запишеться як:

$$(\varnothing \to \{A\})(t) = true \Rightarrow (\varnothing \to \{A\})(t) = true,$$

і виконується тривіально. Отже, таблиця, яка є моделлю множини ФЗ F, знаходиться у НФБК. ▪

■ $|R| = 2$, $|D| \ge 2$ (комірка (3,3)); нехай для визначеності $R = \{A, B\}$ (як в табл. 4.1). Можливі два випадки:

1) нехай множина потенційних ключів – одноелементна, тобто, $\Delta = \{\{A\}\}$ або $\Delta = \{\{B\}\}$, або $\Delta = \{\{A, B\}\}$. Враховуючи доведення з твердження 4.2 для даного випадку, залишається перевірити виконання умови (3) для нетривіальної ФЗ $\varnothing \to \{A\}$ ($\varnothing \to \{B\}$), де атрибут A (атрибут B) є первинним:

$$(\varnothing \to \{A\})(t) = true \Rightarrow (\varnothing \to \{A, B\})(t) = true.$$

Якщо посилка є хибною, то імплікація тривіально істинна. Якщо посилка є істинною, то всі рядки таблиці t мають однакове значення атрибута A. Оскільки $\{A\}$ – потенційний ключ, то у всіх рядків таблиці t однакові значення атрибута B. Отже, імплікація також істинна;

2) нехай тепер множина потенційних ключів двоелементна, тобто, $\Delta = \{\{A\},\{B\}\}$; враховуючи попередній пункт доведення залишається перевірити виконання умови (3) для нетривіальної ФЗ $\{A\} \to \{B\}$ $(\{B\} \to \{A\})$:

$$(\{A\} \to \{B\})(t) = true \Rightarrow (\{A\} \to \{A, B\})(t) = true\,.$$

Імплікація виконується, оскільки за припущенням множина $\{A\}$ є потенційним ключем. Отже, таблиця, яка є моделлю множини ФЗ F, знаходиться у НФБК. ▪□

Змістовна інтерпретація останнього твердження: не більше ніж двоатрибутні таблиці або не більше ніж однорядкові таблиці знаходяться у НФБК (за умови, що вони є моделями множини ФЗ F).

Наслідок 4.4. $НФБК \Leftrightarrow 3НФ$ за умови $|R| \le 2$ або $|D| \le 1$.□

Доведення випливає з тверджень 4.2 і 4.3.□

Наслідок 4.5. $НФБК \Leftrightarrow 3НФ \Leftrightarrow 2НФ$ за умови $|R| \le 2$ або $|D| \le 1$.□

Доведення випливає з наслідків 4.3 і 4.4.□

Лема 4.11. Якщо таблиця t знаходиться у НФБК, то вона знаходиться у 3НФ: $НФБК \Rightarrow 3НФ$. □

Доведення. Нехай таблиця t, яка є моделлю множини ФЗ F, знаходиться у НФБК. З урахуванням включення $R \setminus (\bigcup_{K \in \Delta} K \cup X) \subseteq R \setminus X$ з виконання умови (3):

$$\forall X, \forall A\, (X \subset R \wedge A \in R \setminus X \ \wedge (X \to \{A\})(t) = true \Rightarrow (X \to R)(t) = true)$$

тривіально отримуємо виконання умови (2) для таблиці t:

$$\forall X, \forall A\, (X \subset R \wedge A \in R \setminus (\bigcup_{K \in \Delta} K \cup X) \wedge$$

$$(X \to \{A\})(t) = true \Rightarrow (X \to R)(t) = true)\,.$$

Отже, таблиця t, яка є моделлю множини ФЗ F, знаходиться у 3НФ.□

Наслідок 4.6. Виконуються імплікації: $НФБК \Rightarrow 3НФ \Rightarrow 2НФ$.□

Доведення випливає з лем 4.11 і 4.10. □

4.5 Четверта нормальна форма

Нехай задані множини ФЗ F і БЗЗ G.

Означення 4.10. Скажемо, що таблиця $t(R)$ знаходиться у четвертій НФ (4НФ), якщо, по-перше, таблиця $t(R)$ знаходиться у 1НФ, по-друге, таблиця t є моделлю об’єднання множин ФЗ і БЗЗ $F \bigcup G$ та, по-третє, для кожної нетривіальної БЗЗ $X \rightarrow\rightarrow Y$, такої, що $F \bigcup G \vDash X \rightarrow\rightarrow Y$, множина атрибутів X є суперключем:

$$\forall X, \forall Y (X \subset R \wedge Y \neq R \setminus X \wedge Y \not\subseteq X \wedge$$

$$\wedge (X \rightarrow\rightarrow Y)(t) = true \Rightarrow (X \rightarrow R)(t) = true). \ \square \qquad (4)$$

Твердження 4.4 (достатні умови для знаходження у 4НФ для спеціальних випадків). Таблиця знаходиться у 4НФ якщо $|R| \leq 1$ або $|D| \leq 1$. □

Доведення. Враховуючи результати з табл. 4.1, розглянемо усі можливі випадки для умов $|R| \leq 1$ (перші два стовпчики) або $|D| \leq 1$ (перші два рядки).

■ $|R| = 0$ або $|D| \leq 1$ (комірки (1,1), (2,1), (3,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4)); тобто перший стовпчик і перші два рядки); згідно з табл. 4.1 єдиним потенційним ключем є $\varnothing$, тобто $\Delta = \{\varnothing\}$, звідки елементи множини $2^R = \{X \mid X \subseteq R\}$ є суперключами. Отже, умови 4НФ виконуються автоматично і таблиця, яка є моделлю множини $F \bigcup G$, знаходиться у 4НФ; ▪

■ $|R| = 1$ і $|D| \geq 2$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 4.1). Можливими БЗЗ є $\varnothing \rightarrow\rightarrow \varnothing$, $\varnothing \rightarrow\rightarrow \{A\}$, $\{A\} \rightarrow\rightarrow \varnothing$, $\{A\} \rightarrow\rightarrow \{A\}$, кожна з яких є тривіальною. Отже, умови 4НФ виконуються автоматично (відповідна імплікація тривіально істинна) і таблиця, яка є моделлю множини $F \bigcup G$, знаходиться у 4НФ. ▪□

Зауважимо, що за умови $|R| = 2$ таблиця буде знаходитись у НФБК, але може не знаходитись в 4НФ, як показано в наступному прикладі.

Приклад 4.1. Нехай задані схема $R=\{A,B\}$ і множина БЗЗ $G=\{\varnothing\to\to\{A\}\}$. З вказаних умов випливає, що потенційним ключем є множина атрибутів $\{A,B\}$. Таблиця із схемою $\{A,B\}$, зображена на рис. 4.1, є моделлю заданої множини БЗЗ G. Очевидно, що таблиця знаходиться в НФБК, але, оскільки ліва частина БЗЗ $\varnothing\to\to\{A\}$ не є суперключем, таблиця не знаходиться в 4НФ.

A	B
1	1
1	2

Рисунок 4.1. Таблиця з прикладу 4.1.

Лема 4.12. *Якщо таблиця* $t(R)$ знаходиться у 4НФ, то вона знаходиться у НФБК: $4НФ \Rightarrow НФБК$.□

Доведення. З урахуванням тверджень 4.3 і 4.4 достатньо розглянути випадок, коли $|D|\geq 2$ і $|R|\geq 2$. Нехай таблиця t, яка є моделлю множини $F\bigcup G$, знаходиться у 4НФ і нехай $X\to\{A\}$ – довільна нетривіальна ФЗ (тобто, $A\in R\setminus X$) така, що $F\bigcup G\vDash X\to\{A\}$. Покажемо, що множина атрибутів X є суперключем. Розглянемо випадки:

▪ $X\bigcup\{A\}\subset R$; оскільки, за припущенням, $F\bigcup G\vDash X\to\{A\}$, то за критерієм повноти аксіоматики БЗЗ і ФЗ (теорема 3.3) для випадку $|R|\geq 2$ маємо: $F\bigcup G\vdash X\to\{A\}$; тоді за правилом розширення ФЗ до БЗЗ (лема 3.11, розділ 3) на таблиці t виконується БЗЗ $X\to\to\{A\}$, яка є нетривіальною (нагадаємо, що $A\in R\setminus X$ і $X\bigcup\{A\}\subset R$). За припущенням таблиця t знаходиться у 4НФ, тобто виконується умова (4) і множина атрибутів X є суперключем. Отже, імплікація з умови (3)

$$\forall X\,,\forall A\,(X\subset R\wedge A\in R\setminus X\wedge(X\to\{A\})(t)=true\Rightarrow(X\to R)(t)=true)$$

є істинною і таблиця t знаходиться у НФБК; ▪

▪ $X\bigcup\{A\}=R$; аналогічно до попереднього пункту для випадку, що розглядається ($|R|\geq 2$) маємо: $F\bigcup G\vdash X\to\{A\}$; тоді за правилом поповнення (лема 2.2, розділ 2) на таблиці t виконується ФЗ $X \rightarrow X \bigcup \{A\}$ тобто, ФЗ $X \rightarrow R$, звідси множина атрибутів X є суперключем. Отже, таблиця t, яка є моделлю множини $F \bigcup G$, знаходиться у НФБК. ▪□

Наслідок 4.7. Виконуються імплікації: $4НФ \Rightarrow НФБК \Rightarrow 3НФ \Rightarrow 2НФ$. □

Доведення тривіально випливає з леми 4.12 і наслідку 4.6. □

4.6 Основні результати розділу

В четвертому розділі побудовано фрагмент теорії нормалізації щодо нормальних форм 2-4 порядків в табличних БД.

Формулювання основних понять, а саме, суперключа, потенційного ключа, повної та транзитивної ФЗ, 2-4НФ, виконано з математичною строгістю.

Встановлено значення потенційних ключів у спеціальних випадках, коли значення потужностей множин $|R| \leq 1$ або $|D| \leq 1$.

Розглянуті достатні умови для знаходження таблиці $t(R)$ у кожній з 2НФ-4НФ. Доведено, що в спеціальних випадках, коли $|R| \leq 2$ або $|D| \leq 1$, таблиця $t(R)$, яка знаходиться у 1НФ і є моделлю множини ФЗ F, знаходиться у 2НФ, 3НФ і НФБК.

Показано, що сформульовані означення 2-4 НФ задовольняють принципу «включення» НФ вищих порядків в НФ нижчих порядків, тобто виконуються імплікації: $4НФ \Rightarrow НФБК \Rightarrow 3НФ \Rightarrow 2НФ$.

Основні результати четвертого розділу представлені в табл. 4.2.

Таблиця 4.2

Основні результати розділу 4

№ п.п.	Твердження	Стисле формулювання	Інтерпретація
(1)	(2)	(3)	(4)
1.	Лема 4.1.	Для довільної множини ФЗ F, заданої над схемою R, існує не менше одного потенційного ключа.	Про існування потенційного ключа.

(1)	(2)	(3)	(4)
2.	Результат з підрозділу 4.1.	Кожний мінімальний елемент множини $\langle \Phi_{R\to R} \bigcup \{R \to R\} \bigcap [F], \trianglelefteq \rangle$ є потенційним ключем.	Альтернативне визначення потенційного ключа.
3.	Твердження 4.1.	Таблиця t знаходиться у 2НФ якщо $\mid R \mid \leq 2$ або $\mid D \mid \leq 1$.	Достатні умови для знаходження у 2НФ для спеціальних випадків.
4.	Твердження 4.2.	Таблиця знаходиться у 3НФ якщо $\mid R \mid \leq 2$ або $\mid D \mid \leq 1$.	Достатні умови для знаходження у 3НФ для спеціальних випадків.
5.	Твердження 4.3.	Таблиця знаходиться у НФБК якщо $\mid R \mid \leq 2$ або $\mid D \mid \leq 1$.	Достатні умови для знаходження у НФБК для спеціальних випадків.
6.	Твердження 4.4.	Таблиця знаходиться у 4НФ якщо $\mid R \mid \leq 1$ або $\mid D \mid \leq 1$.	Достатні умови для знаходження у 4НФ для спеціальних випадків.
7.	Наслідок 4.7.	$4\textit{НФ} \Rightarrow \textit{НФБК} \Rightarrow$ $\Rightarrow 3\textit{НФ} \Rightarrow 2\textit{НФ}$	Якщо таблиця знаходиться у НФ «вищого порядку», то вона знаходиться у НФ «нижчого порядку» (для сформульованих у даному розділі означень).

ВИСНОВКИ

Основні результати дисертаційної роботи.

1. Математично доведені повнота та коректність трьох основних аксіоматик залежностей в реляційних БД: по-перше, аксіоматики ФЗ Армстронга, по-друге, аксіоматики БЗЗ, та по-третє, аксіоматики ФЗ і БЗЗ через збіжність відношень синтаксичного та семантичного слідувань.
2. Встановлено критерії повноти аксіоматики ФЗ та аксіоматик ФЗ і БЗЗ в термінах потужностей множини атрибутів та універсального домена.
3. Доведена незалежність складових аксіоматики ФЗ Армстронга (видалення кожної складової зменшує породжуючу силу аксіоматики).
4. Побудована алгебра ФЗ, сигнатура якої складена відповідно до складових аксіоматики ФЗ Армстронга. Розглянута підалгебра тривіальних ФЗ та множина, яка її породжує. Це дозволяє формулювати результати щодо властивостей аксіоматики Армстронга на алгебраїчній мові.
5. Побудовано цілісний та несцперечливий фрагмент математичної теорії нормалізації щодо нормальних форм 2-4 порядків. Для спеціальних випадків встановлені значення потенційних ключів, а також, достатні умови для знаходження таблиці у НФ 2-4 порядків. Сформульовані означення 2-4 НФ задовольняють принципу «включення» НФ вищих порядків в НФ нижчих порядків.

Зазначені результати дозволяють будувати математично обгрунтовані алгоритми нормалізації.

Аналіз отриманих результатів дозволяє зробити наступні висновки.

1. Аксіоматика ФЗ Армстронга є повною, коли в інтерпретації $|D| \geq 2$ або $|R| = 0$.

2. Аксіоматика Армстронга є незалежною в тому розумінні, що без втрати повноти не можна опустити ні єдину аксіому, ні жодне з правил виведення.
3. Аксіоматика БЗЗ є повною, коли в інтерпретації $|R| \leq 1$ або $|R| \geq 2$ і при цьому $|D| \geq 2$.
4. Аксіоматика ФЗ і БЗЗ є повною, коли в інтерпретації $|D| \geq 2$ або $|R| = 0$.
5. НФ знаходяться у логічному зв'язку вигляду 4НФ$\Rightarrow$НФБК$\Rightarrow$ $\Rightarrow$3НФ$\Rightarrow$2НФ.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Aho A. V. The theory of joins in relational databases / A. V. Aho, C. Beeri, J. D. Ullman // Proceedings of the 18th Symp. on Foundations of Computer Science, Providence, R.I., 1977. – P. 107-113.
2. Armstrong W. W. Dependency structures of data base relationships / W. W. Armstrong // Proceedings of the IFIP‘74. North Holland Pub. Co., Amsterdam, 1974. – P. 580-583.
3. Bahar Ö. Normalization and Lossless Join Decomposition of Similarity-Based Fuzzy Relational Databases / Ö. Bahar, A. Yazıcı // International Journal of Intelligent Systems. – 2004. – Vol. 19. – P. 885-917. Published online in Wiley InterScience (www.interscience.wiley.com).
4. Bahmani A. Automatic database normalization and primary key generation / A. Bahmani, M. Naghibzadeh, B. Bahmani // CCECE/CCGEI (Niagara Falls, Canada, May 5-7, 2008). – P. 11-16.
5. Beeri C. A complete axiomatization for functional and multivalued dependencies / C. Beeri, R. Fagin, J. Howard // Proceedings of the ACM-SIGMOD Conf. (Toronto, Canada, Aug. 3-5, 1977). – ACM: New York, 1977. – P. 47-61.
6. Beeri C. A sophisticate’s introduction to database normalization theory / C. Beeri, P. Bernstein, N. Goodman // Proceedings of 4th International Conference on Very Large Data Bases, West Berlin, 1978. – P. 113-124.
7. Beeri C. Computational problems related to the design of normal form relation schemas / C. Beeri, P. A. Bernstein // ACM Transactions on Database Systems. – 1979. – Vol. 4. – № 1. – P. 30-59.
8. Bernstein P. A. Synthesizing Third Normal Form relations from functional dependencies / P. A. Bernstein // ACM Transactions on Database Systems. – 1976. – Vol. 1. – № 4. – P. 277-298.

9. Biskup J. Inferences of multivalued dependencies in fixed and undetermined universes / J. Biskup // Theoretical Computer Science. – 1980. – Vol. 10. – № 1. – P. 93-105.

10. Bui D. Axiomatics for multivalued dependencies in table databases: correctness and completeness / D. Bui, A. Puzikova // Proceedings of the Workshop on Foundations of Informatics – FOI-2015 (August 24-29, 2015, Chisinau, Republic of Moldova). – P. 361-376.

11. Bui D. Axiomatics for Multivalued Dependencies in Table Databases: Correctness, Completeness, Completeness Criteria / D. Bui, A. Puzikova // Theory and Engineering of Complex Systems and Dependability. Series “Advances in Intelligent Systems and Computing”. Vol. 365. – Springer International Publishing Switzerland, 2015. – P. 45-55.

12. Buy D. B. Completeness of Armstrong’s axiomatic / D. B. Buy, A. V. Puzikova // First International Workshop “Critical infrastructure safety and security” – CrISS-Dessert'11 (Kirovograd, Ukraine, May 11-13, 2011). – Kharkiv, KhAI, 2011. Vol. 1(1). – P. 211-215.

13. Buy D. B. Theory of Normalization in Relation Databases (Survey) / D. B. Buy, A. V. Puzikova // Presented at 7th International Conference “Dependable Systems, Services and Technologies DESSERT 2014” (Kiev, Ukraine, May 16-18, 2014). Опубліковано в журналі: Радіоелектронні і комп’ютерні системи №5(69). – Харків “ХАІ”, 2014. – С. 45-49.

14. Casanova M. A. Inclusion dependencies and their interaction with functional dependencies / M. A. Casanova, R. Fagin, C. H. Papadimitriou // Journal of Computer and System Sciences. – 1984. – № 28. – P. 29-59.

15. Chen G. Normalization based on ffd in a fuzzy relational data model / G. Chen, E. E. Kerre, J. Vandenbulcke // Inform Syst, 1996. – Vol. 21. – P. 299-310.

16. Codd E.F. A Relational Model of Data for Large Shared Data Banks / E. F. Codd // Communication of ACM. – 1970.– Vol. 13. – № 6. – P. 377-387.

17. Codd E. F. Further Normalization of the Data Base Relational Model / E. F. Codd // (Presented at Courant Computer Science Symposia Series 6, "Data Base Systems", New York City, May 24-25, 1971.) IBM Research Report RJ909 (August 31, 1971). Republished in Randall J. Rustin (ed.), Data Base Systems: Courant Computer Science Symposia Series 6: Prentice-Hall, 1972.
18. Codd E. F. Recent Investigations into Relational Data Base Systems / E. F. Codd // Proceedings of IFIP Congress'74 (Stockholm, Sweden, August 5-10, 1974). – North-Holland, 1974. – P. 1017-1021.
19. Darwen H. A Normal Form for Preventing Redundant Tuples in Relational Databases / H. Darwen, C. Date, R. Fagin // Proceedings of the 15th International Conference on Database Theory – ICDT'2012 (Berlin, Germany, March 26–30, 2012). – P. 114-126.
20. Date C. J. Database in Depth: Relational Theory for Practitioners / C. J. Date. – California: O'Reilly, 2005. – 208 p.
21. Date C. J. Simple Conditions for Guaranteeing Higher Normal Forms in Relational Databases / C. J. Date, R. Fagin // ACM Transactions on Database Systems. – 1992. – Vol. 17. – № 3. – P. 465-476.
22. Date C. J. Temporal Data and the Relational Model / C. J. Date, H. Darwen, N. Lorentzos. – Morgan Kaufmann, 2002. – 422 p.
23. Dayal U. The fragmentation problem: lossless decomposition of relations into files / U. Dayal, P. A. Bernstein // Proceedings of the ACM SIGMOD international conference on Management of data, 1979. – P. 143-151.
24. Delobel C. Decomposition of a database and the theory of Boolean switching function / C. Delobel, R. Gasey // IBM Journal of Research and Development. – 1973. – Vol. 17. – № 5. – P. 374-386.
25. Elmasri R. Fundamental of Database Systems: [3rd Edition] / R. Elmasri, S. Navathe. – Addison-Wesley, 2000. – 893 p.

26. Fagin R. A Normal Form for Relational Databases That Is Based on Domains and Keys / R. Fagin // Communications of the ACM. – 1981. – Vol. 6. – P. 387-415.

27. Fagin R. Multivalued Dependencies and a New Normal Form for Relational Databases / R. Fagin // ACM Transactions on Database Systems. – 1977. – Vol. 2. – № 1. – P. 262-278.

28. Fagin R. Normal Forms and Relational Database Operators / R. Fagin // Proceedings of the ACM SIGMOD International Conference on Management of Data (Boston, Mass., May 30-June 1). ACM, New York, 1979. – P. 153-160.

29. Fagin R. The Theory of Data Dependencies – a Survey / R. Fagin, M. Y. Vardi // In Mathematics of Information Processing, Proc. Symposia in Applied Mathematics. – Vol. 34, American Mathematical Society. – 1986. – P. 19-71.

30. Fotache M. Why Normalization Failed to Become the Ultimate Guide for Database Designers? [Електронний ресурс] / M. Fotache. – Режим доступу: http://ssrn.com/abstract=905060 or http://dx.doi.org/10.2139/ssrn.905060. – May 1, 2006.

31. Heath I. J. Unacceptable File Operations in Relational Database / I. J. Heath // ACM SIGFIDET Workshop on Data Description, Access, and Control. – San Diego, California. – 1971. – P. 19-33.

32. Isloor S. S. An algorithm with logical simplicity for designing third normal form relations data base schema for functional dependencies / S. S. Isloor // Proceedings of International Conference on DBMS – ICMOD'78 (Fast Milano, Italy, 1978). – P. 31-50.

33. Jaeschke G. Remarks on the Algebra of Non First Normal Form Relations / G. Jaeschke, H. J. Schek // Proceedings of the 1st ACM SIGACT-SIGMOD symposium on Principles of database systems (Los Angeles, California, 1982). – P. 124-138.

34. Levene M. Justification for Inclusion Dependency Normal Form / M. Levene, M. W. Vincent // IEEE Transactions on Knowledge and Data Engineering, 2000. – Vol. 12. – № 2. – P. 281-291.
35. Lin W. Y. Efficient algorithm for BCNF-decomposition / W.-Y. Lin // Information and Software Technology. – 1992. – Vol. 34. – № 5. – P. 308-312.
36. Ling T. W. An Improved Third Normal Form for Relational Databases / T. W. Ling, F. W. Tompa, T. Kameda // ACM Transactions on Database Systems. – 1981. – Vol. 6. – № 2. – P. 329-346.
37. Ling T. W. Logical Database Design with Inclusion Dependencies / T. W. Ling, C. H. Goh // In Proceedings of the Eighth International Conference on Data Engineering, Tempe, Arizona, 1992. – P. 642-649.
38. Makinouchi A. A Consideration of Normal Form on Not-necessarily Normalized Relations in the Relational Data Model / A. Makinouchi // Proceedings of the Third International Conference on Very Large Data Bases (Tokyo, Japan, October 6-8, 1977). – IEEE Computer Society. – 1977. – P. 447-453.
39. Mendelzon A. O. On axiomatizing multivalued dependencies in relational databases / A. O. Mendelzon // ACM Transactions on Database Systems. – 1979. – Vol. 26. – № 1. – P. 37-44.
40. Nicolas J. M. Mutual dependencies and same results on indecomposable relations / J. M. Nicolas // Proceedings of the fourth international conference on Very Large Data Bases, 1978. – Vol. 4. – P. 360-367.
41. Normann R. Minimal lossless decompositions and some normal forms between 4NF and PJ/NF / R. Normann // Information Systems. – 1998. – Vol. 23. – № 7. – P. 509-516.
42. Rissanen J. Independent components of relations / J. Rissanen // ACM Transactions on Database Systems. –1977. – Vol. 2. – № 4. – P. 317-325.

43. Vincent M. W. A corrected 5NF definition for relational database design / M. W. Vincent // Theoretical Computer Science (TCS). – 1997. – Vol. 185.– № 2. – P. 379-391.
44. Vincent M. W. Redundancy Elimination and a New Normal Form for Relational Database Design / M. W. Vincent // In Semantics in Databases (Libkin L., Thalheim B., eds.). – Vol. 1358 of LNCS. – 1998. – P. 247-264.
45. Zaniolo C. Analysis and design of relational schemata for database systems: Ph.D. dissertation, Tech. Rep. UCLA-Eng-7769, Dep. Computer Science, Univ. California at Los Angeles, July 1976.
46. Zaniolo C. A New Normal Form for the Design of Relational Database Schemata / C. Zaniolo // ACM Transactions on Database Systems. – 1982. – Vol. 7. – № 3. – P. 489-499.
47. Биркгоф Г. Теория решеток / Г. Биркгоф. – Москва: Наука, 1984. – 564 с.
48. Редько В. Н. Аксіоматика багатозначних залежностей табличних баз даних / В. Н. Редько, Д. Б. Буй, А. В. Пузікова // Доповіді НАН України. – 2015. – № 6. – С. 24-29.
49. Буй Д. Б. Аксіоматика багатозначних залежностей табличних баз даних: коректність, повнота, критерій повноти / Д. Б. Буй, А. В. Пузікова // Міжнародна наукова конференція «Сучасні проблеми математичного моделювання та обчислювальних методів» (Рівне, Україна, 19-22 лютого 2015 р.). Опубліковано в Волинському математичному віснику. Сер.: прикладна математика. – 2015. – Випуск 11 (20). – С. 172-189.
50. Буй Д. Б. Аксіоматика багатозначних залежностей табличних баз даних: повнота та її критерій / Д. Б. Буй, А. В. Пузікова // Матеріали міжнародної наукової конференції «Теоретичні та прикладні аспекти побудови програмних систем» – TAAPSD'2014 (Україна, Київ, 15-17 грудня 2014 року). – С. 35-43.
51. Буй Д. Б. Деякі некласичні нормальні форми в реляційних базах даних / Д. Б. Буй, А. В. Пузікова // Вісник Київського національного

університету імені Тараса Шевченка. Сер.: фіз.-мат. науки. – 2015. – № 1. – С. 65-74.

52. Буй Д. Б. Критерій повноти аксіоматики Армстронга / Д. Б. Буй, А. В. Пузікова // Матеріали міжнародної конференції «Теоретичні та прикладні аспекти побудови програмних систем» – TAAPSD'2011 (Україна, Ялта, 19-23 вересня 2011 року). – С. 30-34.
53. Буй Д. Б. Критерии полноты аксиоматик зависимостей в табличных базах данных / Д. Б. Буй, А. В. Пузикова // Материалы XVII Международной конференции «Дискретные модели в теории управляющих систем» (Москва и Подмосковье, 20–22 мая 2015 г.). – М.: МАКС Пресс, 2015. – С. 37-39.
54. Буй Д. Б. Математична теорія нормалізації / Д. Б. Буй, А. В. Пузікова // Вісник Київського національного університету імені Тараса Шевченка. Сер.: фіз.-мат. науки. – 2015. – № 2. – С. 103-112.
55. Буй Д. Б. Незалежність аксіоматики Армстронга та алгебра функціональних залежностей / Д. Б. Буй, А. В. Пузікова // Штучний інтелект. – 2015. – № 1-2. – С. 121-126.
56. Буй Д. Б. Незалежність аксіоматики функціональних залежностей Армстронга / Д. Б. Буй, А. В. Пузікова // Матеріали IX Міжнародної конференції «Інтернет-Освіта-Наука-2014» – ІОН-2014 (Вінниця, 14-17 жовтня, 2014). – Вінниця: ВНТУ, 2014. – С. 55-56.
57. Буй Д. Б. Обзор современной теории нормализации в реляционных базах данных / Д. Б. Буй, А. В. Пузикова // Проблемы теоретической кибернетики. Материалы XVII Международной конференции (Казань, 16–20 июня 2014 г.). Под редакцией Ю.И. Журавлева. – Казань: Отечество, 2014. – С. 39-43.
58. Буй Д. Б. Огляд теорії нормалізації в реляційних базах даних / Д. Б. Буй, А. В. Пузікова // Труды международной научно-технической конференции «Компьютерное моделирование в наукоемких технологиях» (Харьков, 28-31 мая 2014 г.). – С. 60-63.

59. Буй Д. Б. Повнота аксіоматики Армстронга / Д. Б. Буй, А. В. Пузікова // Вісник Київського національного університету імені Тараса Шевченка. Сер.: фіз.-мат. науки. – 2011. – № 3. – С. 103-108.

60. Буй Д. Б. Полнота аксиоматики Армстронга / Д. Б. Буй, А. В. Пузикова // Материалы XVI Международной конференции «Проблемы теоретической кибернетики» (Нижний Новгород, 20-25 июня, 2011). – Изд-во Нижегородского госуниверситета, 2011. – С. 85-88.

61. Буй Д. Б. Теорія нормалізації в реляційних базах даних: сучасний стан / Д. Б. Буй, А. В. Пузікова // Наукові записки НаУКМА. Сер.: комп‘ютерні науки. – 2015. – Т. 177. – С. 83-92.

62. Буй Д. Б. Теорія нормалізації табличних баз даних: 2-3НФ, НФБК / Д. Б. Буй, А. В. Пузікова // Матеріали X Міжнародної конференції «Теоретичні та прикладні аспекти побудови програмних систем» – TAAPSD'2013 (Ялта, 25 травня – 2 червня, 2013). – С. 34-40.

63. Буй Д.Б. Властивості теоретико-множинних конструкцій повного образу та обмеження / Д.Б. Буй, Н.Д. Кахута // Вісник Київського університету. Сер.: фіз.-мат. науки. – 2005. – Вип. 2. – С. 232-240.

64. Виноградова М. В. Конструктор баз данных на основе сущностей и их реквизитов с возможностью нормализации [Электронный ресурс] / М. В. Виноградова, Э. Г. Игушев // Электронное научно-техническое издание «Наука и образование», 2011. – № 10. – Режим доступа: http://technomag.edu.ru/doc/242645.html.

65. Григорьев Ю. А. Алгоритм синтеза частично оптимальной схемы реляционной базы данных [Электронный ресурс] / Ю. А. Григорьев // Электронное научно-техническое издание «Наука и образование», 2012. – № 1. – Режим доступа: http://technomag.edu.ru/doc/294486.html.

66. Дейт К. Дж. Введение в системы баз данных / К. Дж. Дейт. – М.: Наука, 1980. – 464 с.

67. Дейт К.Дж. Введение в системы баз данных: [8-е изд.: пер. с англ.] / К.Дж. Дейт. – Москва: Издательский дом "Вильямс", 2005. – 1328 с.

68. Дрибас В. П. Реляционные модели баз данных / В. П. Дрибас. – Минск: Изд-во БГУ им. В. И. Ленина, 1982. – 192 с.
69. Зорин И. Теоретико-графовое приведение реляционной базы данных к третьей нормальной форме Э. Кодда [Електронный ресурс] / И. Зорин // Электронное научное издание «Устойчивое инновационное развитие: проектирование и управление», 2009. – Т. 5. – С. 50-59. – Режим доступа: www.rypravlenie.ru.
70. Исаченко А. Н. Модели данных и СУБД / А. Н. Исаченко, С. П. Бондаренко. – Минск: БГУ, 2007. – 205 с.
71. Линдон Р. Заметки по логике / Р. Линдон. – Москва: Мир, 1968. – 128 с.
72. Мальцев А.И. Алгоритмы и рекурсивные функции / А.И. Мальцев. – М.: Наука, 1986. – 368 с.
73. Мейер Д. Теория реляционных баз данных: [пер. с англ.] / Д. Мейер. – Москва: Мир, 1987. – 608 с.
74. Неклюдова Е. А. Синтез логической схемы реляционной базы данных / Е. А. Неклюдова, М. Ш. Цаленко // Программирование. – 1979. – № 6. – С. 58-68.
75. Петров С. В. Об аксиоматизации зависимостей по соединению / С. В. Петров // Применение методов математической логики: Тезисы докл. IV Всес. конф. «Представление знаний и синтез программ». – Таллин: АН ЭССР, 1986. – С. 151-152.
76. Реляційні бази даних: табличні алгебри та SQL-подібні мови / В.Н. Редько, Ю. Й. Брона, Д. Б. Буй, С. А. Поляков. – Київ: Видавничий дім "Академперіодика", 2001. – 198 с.
77. Скорняков Л. А. Элементы теории структур / Л. А. Скорняков. – Москва: Наука, 1982. – 160 с.
78. Ульман Дж. Основы систем баз данных: [пер. с англ.] / Дж. Ульман. – Москва: Финансы и статистика, 1983. – 334 с.
79. Филиппович А. Взаимные функциональные зависимости / А. Филиппович // Системный администратор. – 2002. – № 1. – С. 84-89.

80. Цаленко М. Ш. Моделирование семантики в базах данных / М. Ш. Цаленко. – М.: Наука, 1989. – 287 с.

81. Шенфилд Дж. Математическая логика / Дж. Шенфилд. – Москва: Наука, 1975. – 528 с.

ДОДАТКИ

Додаток А

Класичні та основні некласичні НФ

№ за п/п	Назва НФ	Автор(и), рік, посилання	Означення
(1)	(2)	(3)	(4)
1.	1НФ	К. Дж. Дейт, 1980, [65]	Відношення знаходиться в 1НФ, якщо кожне значення у відношенні – тобто, кожне значення домену в кожному кортежі – є атомарним (таким, що не розкладається) елементом даних, наприклад, число або рядок символів
2.	1НФ	R. Elmasri, S. R. Navathe, 2000, [25]	Відношення знаходиться в 1НФ, якщо домен кожного атрибута містить тільки атомарні (прості, неподільні) значення, і значення кожного атрибуту набувають тільки повних значень з цього домену
3.	1НФ	К. Дж. Дейт, 2005, [67]	Змінна-відношення (relvar) знаходиться у 1НФ тоді і тільки тоді, коли в кожному її допустимому значенні кожний кортеж містить тільки одне значення для кожного атрибуту
4.	2НФ	Д. Мейер, 1987, [73]	Схема відношення R знаходиться у 2НФ відносно множини ФЗ F, якщо вона знаходиться у 1НФ і кожен непервинний атрибут повністю залежить від кожного ключа для R

Продовж. Додатку А

(1)	(2)	(3)	(4)
5.	3НФ	Д. Мейер, 1987, [73]	Схема відношення R знаходиться у 3НФ відносно множини ФЗ F, якщо вона знаходиться у 1НФ і жоден з непервинних атрибутів в R не є транзитивно залежним від ключа для R
6.	3НФ	К. Дж. Дейт, 2005, [67]	Змінна-відношення знаходиться у 3НФ тоді і тільки тоді, коли вона знаходиться у 2НФ і кожний неключовий атрибут нетранзитивно залежить від її первинного ключа (в означенні передбачається наявність тільки одного потенційного ключа, який і є первинним ключем відношення)
7.	3НФ	В. П. Дрібас, 1982, [68]	3НФ відношення $r(R)$ – це або дане відношення, якщо воно знаходиться у 2НФ і не містить транзитивних залежностей непервинних атрибутів від потенційних ключів, або набір яких-небудь його проекцій, кожна з яких задовольняє вказаним умовам, причому дане відношення повинно відновлюватись через природнє з'єднання цих проекцій

Продовж. Додатку А

(1)	(2)	(3)	(4)
8.	Elementary Key Normal Form (EKNF)	C. Zaniolo, 1982, [46]	Відношення із схемою R знаходиться у EKNF за означенням, якщо для кожної нетривіальної ФЗ $X \to A$: 1) X є суперключем відношення, або 2) атрибут A належить деякому елементарному ключу (ключ X є елементарним, якщо для *деякого* атрибута $A \in R$ ФЗ $X \to A$ є повною)
9.	НФБК	C. Zaniolo, 1982, [46]	Відношення із схемою R знаходиться у НФБК, якщо для кожної нетривіальної ФЗ $X \to A$ множина атрибутів X є суперключем
10.	НФБК	R. Fagin, 1979, [28]	Схема відношення знаходиться у НФБК, якщо вона знаходиться у 1НФ і кожна ФЗ $\sigma \in \Sigma$ є логічним наслідком множини ключів Δ $(\Delta \models \sigma)$
11.	4НФ	R. Fagin, 1977, [27]	Схема відношення знаходиться у 4НФ, якщо вона знаходиться у 1НФ і кожна БЗЗ $\sigma \in \Sigma$ є логічним наслідком множини ключів Δ $(\Delta \models \sigma)$
12.	PJ/NF (класична 5НФ)	R. Fagin, 1979, [28]	Два означення: 1) схема відношення знаходиться у PJ/NF, якщо вона знаходиться у 1НФ і для множини ключів Δ та кожної ЗЗ $\sigma \in \Sigma$ алгоритм приналежності видає результат «*True*»;

Продовж. Додатку А

(1)	(2)	(3)	(4)
			2) схема відношення знаходиться у PJ/NF, якщо вона знаходиться у 1НФ і кожна ЗЗ $\sigma \in \Sigma$ є логічним (семантичним) наслідком множини ключів Δ $(\Delta \models \sigma)$
13.	5НФ	К. Дж. Дейт, 2005, [67]	Змінна-відношення знаходиться в 5НФ (PJ/NF) за означенням, якщо кожен компонент кожної нетривіальної ЗЗ є суперключем
14.	Попередня 5НФ	Д. Мейер, 1987, [73]	Схема відношення R знаходиться у попередній PJ/NF за означенням, якщо кожна ЗЗ $*(R_1, R_2, ..., R_n)$ $(R_1 \cup R_2 \cup ... \cup R_n = R)$, яка виводиться (синтаксично слідує) з множини обмежень Σ, є або тривіальною, або її компоненти є суперключами
15.	Reduced-5NF (5NFR)	M. W. Vincent, 1997, [43]	Нехай Σ – множина ФЗ і ЗЗ, причому кожна ЗЗ $*(R_1, R_2, ..., R_p)$ задовольняє умовам: 1) $R_1 \cup R_2 \cup ... \cup R_p = R$, де R – схема відношення; 2) $\forall i,\ 1 \le i \le p$, $*(R_1, R_2, ..., R_{i-1}, R_{i+1}, ..., R_p)$ або не належить Σ^+, або $R_1 \cup R_2 \cup ... \cup R_{i-1} \cup R_{i+1} \cup ... \cup R_p \neq R$, де Σ^+ – множина усіх залежностей, які слідують (синтаксично) з множини Σ.

Продовж. Додатку А

(1)	(2)	(3)	(4)
			Пара $\langle R,\Sigma\rangle$ знаходиться у 5NFR за означенням, якщо ліва частина кожної її нетривіальної ФЗ є суперключем і для кожної ЗЗ $\sigma\in\Sigma$ кожен компонент σ є суперключем
16.	Superkey Normal Form (SKNF)	R. Normann, 1998, [41]	Реляційна схема $\langle R,\Sigma\rangle$ знаходиться у НФ, яка визначається суперключем (SKNF), якщо кожен компонент кожної нескоротної ЗЗ є суперключем
17.	Key-Complete Normal Form (KCNF)	M. W. Vincent, 1998, [44]	Пара $\langle R, \Sigma\rangle$ знаходиться у KCNF, якщо ліва частина кожної нетривіальної ФЗ $\varphi\in\Sigma$ є суперключем, а кожна ЗЗ $\sigma\in\Sigma$ повністю визначається ключами
18.	Redandancy-free normal form (RFNF)	M. W. Vincent, 1998, [44]	Пара $\langle R, \Sigma\rangle$ знаходиться у надлишково-вільній нормальній формі (RFNF), якщо не існує відношення $r(R)\in SAT(\Sigma)$, яке є надлишковим, де $SAT(\Sigma)$ – множина екземплярів (моделей) множини Σ, яка складається з ФЗ і ЗЗ
19.	Essential Tuple Normal Form (ETNF)	H. Darwen, C. Date, R. Fagin, 2012, [19]	Реляційна схема $\langle R, \Sigma\rangle$ знаходиться в кортеже-необхідній НФ (ETNF), якщо кожний кортеж кожного екземпляра схеми $\langle R, \Sigma\rangle$ не є частково чи повністю надлишковим

Продовж. Додатку А

(1)	(2)	(3)	(4)
20.	Domain-Key Normal Form (DK/NF або ДКНФ)	R. Fagin, 1981, [26]	Відношення знаходиться у DK/NF, якщо кожне обмеження з його схеми є логічним наслідком з об’єднання множин залежності ключа та залежності домена

Додаток Б

Бібліографія з теорії залежностей та нормалізації

№ з/п	Автор(и), рік, посилання	Головні результати
(1)	(2)	(3)
1.	E. Codd, 1970, [16]	Наводяться змістовні означення активного домену, первинного і зовнішнього ключів та процедура усунення непростих доменів, для якої використовується термін – нормалізація
2.	E. Codd, 1971, [17]	Обговорюються цілі подальшої нормалізації відношень, які знаходяться в 1НФ, зокрема, звільнення колекції відношень від залежностей (аномалій) вставки, оновлення та знищення. Розглядається концепція ФЗ, демонструється можливість використання ФЗ для розв’язання проблем проектування БД, наводяться визначення 2НФ, транзитивної ФЗ та 3НФ, наводиться спосіб зведення до вказаних НФ
3.	I. Heath, 1971, [31]	Наводиться теорема про виконуваність декомпозиції без втрат відношення $R(A,B,C)$ на проекції $R_1 = \pi_R(A,B)$ та $R_2 = \pi_R(A,\mathrm{C})$ за умови, якщо виконується ФЗ $A \to B$. Формулюється означення 3НФ, яке насправді є вперше офіційно опублікованим означенням НФБК
4.	C. Delobel, R. Gasey, 1973, [24]	Розглядаються аксіоми виведення для ФЗ. Наводиться алгоритм зведення реляційної схеми до 3НФ, який є неточним – у контрприкладі з [68] показано, що на виході може бути отримана схема, яка не знаходиться у 3НФ

Продовж. Додатку Б

(1)	(2)	(3)
5.	E. Codd, 1974, [18]	Формулюється означення посиленої 3НФ, у якому не використовується поняття транзитивної ФЗ (інша назва – нормальна форма Бойса-Кодда (НФБК))
6.	W. Armstrong, 1974, [2]	Наводиться точна характеристика потенційних ключів. Викладається формальна теорія ФЗ, будується аксіоматика та розглядається її повнота і коректність, пропонується метод побудови відношення Армстронга для множини ФЗ
7.	C. Zaniolo, 1976, [45]	Розглянуто концепцію БЗЗ та приклади їх використання
8.	R. Fagin, 1977, [27]	Досліджуються властивості БЗЗ та розглядається концепція 4НФ
9.	C. Beeri, R. Fagin, J. Howard, 1977, [5]	Наводиться строгий та повний набір правил виведення для БЗЗ, а також правила, які пов’язують ФЗ та БЗЗ. Доводиться повнота системи аксіом для ФЗ і БЗЗ.
10.	J. Rissanen, 1977, [42]	Розглянуто концепцію ЗЗ (у роботі названі взаємозалежностями), обговорюються залежності проекцій, які можуть стати причиною аномалій
11.	A. Aho, C. Beeri, J. Ullman, 1977, [1]	Досліджено ЗЗ, розглянуто поняття з’єднання без втрат та тривіальної ЗЗ як такої, що множина атрибутів хоча б однієї з проекцій співпадає зі схемою відношення; наведено алгоритм перевірки виконуваності декомпозиції без втрат
12.	J. Nicolas, 1978, [40]	Розглянуто вкладені залежності як окремий випадок ЗЗ

Продовж. Додатку Б

(1)	(2)	(3)
13.	U. Dayal, P. Bernstein, 1978, [23]	Досліджено властивості ЗЗ (у роботі названі взаємозалежностями)
14.	A. Mendelzon, 1979, [39]	Обговорюється незалежність системи аксіом для БЗЗ
15.	R. Fagin, 1979, [28]	Представлено концепцію PJ/NF або класичної 5НФ
16.	R. Fagin, 1981, [26]	Представлено концепцію ДКНФ, яка включає в себе усі попередні та не визначається в термінах традиційних залежностей (ФЗ, БЗЗ чи ЗЗ)
17.	T. Ling, F. Tompa, T. Kameda, 1981, [36]	Розглядається концепція покращеної НФ (An Improved Third Normal Form) та пропонується алгоритм, який за заданою множиною ФЗ будує множину відношень, у якій відсутні «зайві» атрибути
18.	C. Zaniolo, 1982, [46]	Сформульовано більш строге означення 3НФ: «Відношення R знаходиться у 3НФ тоді і тільки тоді, коли для кожної нетривіальної ФЗ $X \to A$, (a) X є суперключем, або (b) A є первинним атрибутом», яке надає чітке уявлення про різницю між 3НФ і НФБК (остання не містить умову (b)). Визначається нова нормальна форма EKNF, яка займає проміжне положення між 3НФ та НФБК
19.	С. Петров, 1986, [75]	Отримано результат про відсутність повної скінченної множини правил виведення для ЗЗ
20.	C. Date, R. Fagin, 1992, [21]	Розглянута еквівалентність 3-5НФ в особливих випадках, коли потенційні ключі (ключ) є простими

Продовж. Додатку Б

(1)	(2)	(3)
21.	T. Ling, C. Goh, 1992, [37]	Представлено концепцію вкладеної нормальної форми (Inclusion Normal Form), яка визначається в термінах ФЗ та вкладених залежностей, включає в себе покращену 3НФ [46] та може бути отримана за допомогою поширення алгоритму Deletion Normalization Algorithm [46]
22.	M. Vincent, 1998, [44]	Представлено концепцію нормальної форми з повним ключем (Key-Complete Normal Form), яка визначається в термінах ФЗ та ЗЗ і є більш слабкою, ніж 5НФ
23.	M. Levene, M. Vincent, 2000, [34]	Представлено концепцію нормальної форми із вкладеними залежностями (Inclusion Dependency Normal Form)
24.	C. Date, H. Darwen, N Lorentzos, 2002, [22]	Розглянуто концепцію шостої нормальної форми (6НФ) для хронологічних БД
25.	H. Darwen, C. Date, R. Fagin, 2012, [19]	Розглянуто концепцію нормальної форми з необхідними кортежами (Essential Tuple Normal Form), яка визначається в термінах ФЗ і ЗЗ та займає проміжне положення між 4НФ та 5НФ