***3. Залежності та нормальні форми реляцій.***

***3.1. Функціональні залежності (ФЗ).***

***Квазіключ, ключ, суперключ, «чужі ключі».***

Згадаємо поняття функціональної залежності. Нехай задано бінарне відношення

R D1D2, тоді будемо говорити, що D1 **функціонально визначає** D2 (а D2 **функціонально залежить** від D1), якщо d1 ϵ D1 і |imR d1|= 1.

Розглянемо тепер випадок n-арного відношення

R D1D2…Dn, n≥2; нехай М1, М2 – списки атрибутів відношення R, тоді

r1 ϵ R[M1], r2 ϵ R[M2] розглянемо відношення τR:

r1 τR r2 ↔  r ϵ R , де r1=r[M1] & r2=r[M2]

Таким чином, n-арне відношення зводиться до бінарного, а для бінарного відношення поняття функціональної залежності ми визначили. Отже, якщо τR є функціональним, то ми можемо говорити, що М1 функціонально визначає М2:

R.M1→RM2; а якщо до того ж τR є взаємно-однозначним, то

R.M1 ↔ R.M2;

Якщо ΩR – це множина імен атрибутів деякої реляції R, то ℬ(ΩR) – це булеан, тобто множина всіх підмножин множини ΩR.

Пару (ℬ(ΩR),fR) будемо називати **структурою** функціональних залежностей реляції R, де fR – множина функціональних залежностей для реляції R. Дещо пізніше ми переконаємося, що ця пара дійсно є структурою у строго математичному сенсі.

Розглянемо тепер поняття квазіключа (candidate of key), первинного ключа чи ключа (primary key), суперключа (super-key), «чужого» ключа (foreign key). Поняття ключа відповідно до реляційного підходу є дуже близьким до поняття ключа в ER-моделі, але не тотожне.

Квазіключем К ΩR (деякої реляції R) називається список атрибутів, який задовольняє умовам :

1)М ΩR , R.K→R.M

2) К´ К (власної підмножини), М ΩR: **¬(**R.K´→ R.M)

Інколи в літературі зустрічається дещо інше визначення квазіключа, але воно еквівалентне даному:

замість умови 1) ставлять умову 1΄) R.K→R. ΩR , а замість умови 2) ставлять умову

2΄)  К´ К, К˝= K\К˝: **¬(**R.К´→ R.К˝)

Розглянемо тепер приклад реляції:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| А1 | А2 | А3 | А4 | А5 | А6 | А7 |
| Тріска | 08611 | жива | 200 | 400 | 140 | 60 |
| Тріска | 08611 | заморожена | 30 | 50 | 20 | 10 |
| Тріска | 08611 | консерв. | 100 | 75 | 25 | 75 |
| Судак | 08612 | жива | 100 | 150 | 34 | 66 |
| Судак | 08612 | заморожена | 150 | 200 | 75 | 75 |
| Судак | 08612 | консерв. | 100 | 40 | 35 | 65 |

де А1 – назва продукції, А2 – шифр, А3 – стан, А4 – фактично одержано, А5 – замовлено, А6 – одержано 1-го сорту, А7 – одержано 2-го сорту.

Також для цієї реляції задана така структура функціональних залежностей:

R.A1↔ R.A2; R.(A6,A7) →R.A4; R.(A4,A6) →R.A7; R.(A4,A7) →R.A6

На основі описаної структури функціональних залежностей може бути створено кілька квазіключів. Атрибути А5 і А3 не входять до структури функціональних залежностей, тому А5 і А3 повинні обов’язково входити у кожен квазіключ. З А4,А6, А7 будь-які два повинні ввійти у кожен квазіключ. З A1 і A2 – один ввійде до складу квазіключа. Отже випишемо деякі квазіключі: А1А3А5A4A6, А2А3А5A4A6, А1А3А5A4A7,…

Тепер перед розробником бази даних постає задача – обрати з множини квазіключів один первинний ключ; на цей вибір вже не впливають функціональні залежності, але це не означає, що він повністю довільний. Інші причини стають вагомими. Наприклад, по А1 розмір поля змінюється в широкому діапазоні, а по А2 розмір фіксований, отже з точки зору ефективності обробки А2 має перевагу. З іншого боку, А1 має вищу семантичну навантаженість. Важливими факторами також можуть бути стійкість до помилок певних атрибутів (їх доменів) при ручному вводі даних.

***3.2. Первинні та вторинні атрибути. 1-а нормальна форма (1НФ); Функціонально повна залежність (ФПЗ), 2НФ, 2НФп; Теорема Хіза.***

Атрибути, які входять до складу хоча б одного квазіключа називаються **первинними**.

Атрибути, які не входять до складу жодного квазіключа називаються **вторинними**.

Кажуть, що реляція знаходиться в першій нормальній формі (**1НФ**), якщо всі її атрибути атомарні (тобто неподільні). Оскільки всі таблиці, з якими ми мали справу до цього часу, були в 1НФ, то поговоримо про реляції не в 1НФ. Приклад: відомість на заробітну плату. Деякі стовпчики цієї таблиці є складеними (наприклад, «Нараховано», «Утримано»), тобто в їх складі в свою чергу знаходяться інші стовпчики. Процес перетворення такої таблиці до 1НФ називається нормалізацією до 1НФ, або просто **нормалізацією**.

М2 **функціонально повно залежить** (ФПЗ) від М1, якщо

1. R.M1 → R.M2

2.АМ1 (власної підмножини М1) ВМ2: **¬(**R.А→R.В)

Кажуть, що реляція знаходиться в 2-ій нормальній формі (**2НФ**), якщо вона знаходиться в першій нормальній формі і кожний вторинний атрибут функціонально повно залежить від кожного квазіключа.

На перший погляд може скластися враження, що наведене визначення дуже далеке від практичних потреб, але це не так. Розглянемо приклад таблиці Т:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Т | КП | КД | місто | КП→ місто – це єдина залежність. |
|  |  |  |  |

Квазіключем і ключем є (КП, КД), а місто - вторинний атрибут. Маємо залежність від КП і КД. Оскільки є залежність КП→ місто, то реляція не знаходиться в 2НФ. А тепер проблеми з практики: нехай в таблиці знаходиться інформація, що постачальник П1 з міста О постачає деталь Д1. В якийсь момент він припиняє поставку, тоді треба вилучати відповідний кортеж, але якщо це був єдиний кортеж з постачальником П1, то ми втратимо інформацію про те, що П1 живе в місті О. Через деякий час П1 відновив поставки деталі Д7 і відповідний кортеж має бути занесений в таблицю, але в ньому немає даних про місто постачальника П1. Проблеми такого типу були відомі задовго до появи реляційного підходу, вирішувались інтуїтивним способом і називались відповідно **аномаліями** вилучення та вставки. Була також аналогічна проблема – аномалія оновлення.

Вирішення цих проблем полягає у декомпозиції реляції Т на Т1=Т[КП,місто] і Т2=Т[КД,КП]; у такому випадку згадані проблеми зникають. Відзначимо також, що реляція Т може бути відновлена з 2-х реляцій без втрати інформації за допомогою операції природного з’єднання по атрибуту КП. Дійсно, оскільки КП і КД утворюють ключ реляції Т, то в Т2 ми будемо мати точно таку ж кількість кортежів, що і в початковій реляції Т (вилучення дублів неможливе), а доповнення значеннями міста (3-й атрибут) здійснюється з Т1 однозначно, бо має місце функціональна залежність КП→ місто. Реляції Т1 і Т2 знаходяться в 2НФ, оскільки ключ з Т1 складається з одного атрибута, а реляція Т2 не має вторинних атрибутів.

Достатні умови знаходження реляції в 2НФ.

1. всі атрибути первинні;
2. кожен квазіключ має один атрибут.

Теорема Heath I.Y.:

Якщо R.M1 → R.M2, тоді R можна декомпонувати на дві реляції з можливістю відновлення без втрат. Природне з’єднання відбувається по атрибутам М1.

R =(R[M1,M2]) [M1 M2] R[ΩR \(M2\M1)]

ΩR \(M2\M1) = M1 **¬(**M2)

Ідея доведення – по суті та ж сама, що і у попередньому прикладі.

Беремо реляцію R[ΩR\(M2\M1)] (вона має ту ж саму кількість кортежів, що і R) і доповнюємо її елементами з М2, які однозначно визначаються на основі функціональної залежності від М1, в результаті отримуємо повну реляцію R.

Процедура отримання реляцій у 2НФ наступна: потрібно дослідити залежності заданої реляції і визначити, чи є порушення 2НФ. Якщо порушень немає, то процедура завершена і ми маємо 2НФ. Якщо ж порушення є, то потрібно взяти залежність, яка створює це порушення, і застосувати теорему Хіза. В результаті отримаємо дві реляції, стосовно яких знову застосуємо попередню процедуру. Оскільки кількість як залежностей, так і атрибутів у початковій реляції скінченна, то рано чи пізно вийдемо на 2НФ. Цей процес називається декомпозицією реляції до 2НФ, або нормалізацією (2НФ). Основна проблема полягає в тому, що таких порушуючих 2НФ залежностей може бути кілька, а вибір тої чи іншої призводить до різних кінцевих результатів.

**Оптимальною декомпозицією** вважається та, у якої найменше реляцій. Тому для отримання оптимальної декомпозиції потрібно отримати всі, а потім вибрати серед них оптимальну. Очевидно, що на практиці застосовувати алгоритми переборного типу не є достатньо ефективним рішенням, тому використовують методики, про які мова піде пізніше.

Кажуть, що реляція знаходиться в 2-ій нормальній формі посиленій(**2НФп**)(або у 2-ій нормальній формі *Бойса-Кодда* (Boyce-Codd)), якщо вона знаходиться в першій нормальній формі і **кожен** атрибут функціонально повно залежить від кожного квазіключа. Іншими словами знімається вимога вторинності атрибута.

***3.3. Транзитивна функціональна залежність в сенсі реляційного підходу, 3НФ, 3НФп чи нормальна форма Бойса-Кодда. Декомпозиція реляцій до 3НФ.***

Нехай М1 і М2, М3 ΩR, М1≠М2, М3 М2

М3 **транзитивно** залежить від М1, якщо R.M1  R.M2 & R.M2  R.M3 & **¬**(R.M2  R.M1).

Відзначимо, що **¬**(R.M3  R.M1) є наслідком.

Зобразимо це графічно:

Зауважимо, що транзитивна залежність у сенсі реляційного підходу відрізняється від поняття транзитивної залежності у класичному сенсі тим, що вимагається

**¬**(R.M2  R.M1).

Кажуть, що реляція знаходиться в **3НФ**, якщо вона в 2 НФ і не має транзитивної залежності вторинних атрибутів від кожного квазіключа.

Кажуть, що реляція знаходиться в **3НФп**, якщо вона в 2 НФп і не має має транзитивної залежності кожного атрибута від кожного квазіключа.

А5 А4 А3

Приклад: А1 – шифр міністерства,

А2 – шифр головного управління,

А3 – шифр області,

А6  А2 А4 – шифр району,

А6 – шифр підприємства,

А7 – шифр галузі.

А1

{ А5А6А1}

{ А5А2А1} А5 є ключем ієрархічної структури.

{ А5А4А3}

Маємо транзитивну залежність А5А4А3, причому це транзитивна залежність як в класичному, так і в реляційному сенсі. Залежності А5А6А1 як і А5А2А1 -також транзитивні в обох сенсах. А от залежність А6А2С (чи А2А6А1) є транзитивною в класичному сенсі, але не транзитивною в реляційному. Такі структури в логічному проектуванні називають «трикутником» (жаргонна назва).

В даному прикладі існують 2 транзитивні залежності і тільки один первинний атрибут, тому ця реляція не знаходиться в 3 НФ. Для декомпозиції заданої реляції до 3НФ скористаємося теоремою Хіза. Відсікаємо А4А3. У відповідністю з теоремою Хіза отримаємо R1(А4А3) i R1(А5А6А1,А2,А4). Отримали 2 реляції: R1 знаходиться в 3НФ, а R1 – ні, продовжуємо декомпонувати R1, наприклад, по залежності А5А6А1. Отримуємо: R2(А5А6) і R2(А6А2,А1) – обидві знаходяться в 3НФ. Отже результат декомпозиції R1(А4А3); R2(А5А6); R3(А6А2,А1) = R2. Методика, яка була задіяна для цього прикладу, характерна для ієрархічних структур: за допомогою теореми Хіза потрібно відсікати листові вершини і рухатися вгору до кореня.

Достатньою умовою знаходження реляції в 3НФ є структура типу “сонечко”: в центрі знаходяться ключові атрибути, а на променях, довжиною в одну залежність – вторинні атрибути.

***3.4. Багатозначна залежність. 4НФ. Залежність по з’єднанню без втрат. 5НФ.***

4НФ базується не на функціональних, а на багатозначних залежностях, тому спочатку поговоримо про них.

|  |  |  |  |
| --- | --- | --- | --- |
| КВП | Курс | Викладачі | Підручник |
|  | програмування | Іванчук | Pascal |
|  | програмування | Іванчук | C |
|  | програмування | Сидоренко | Pascal |
|  | програмування | Сидоренко | C |

Розглянемо таблицю КВП. Курс програмування ведуть два викладачі і для курсу використовується два підручники. Легко бачити, що функціональних залежностей між атрибутами цієї таблиці немає; разом з тим якісь залежності є – певний курс можуть вести не всі викладачі, а тільки деяка їх підмножина, подібні ж стосунки між курсом та підручниками, а також викладачами та підручниками. Залежності такого виду будемо називати багатозначними (multivalued).

Перейдемо до формального визначення поняття багатозначної залежності.

Нехай Х і У списки атрибутів реляції R.

Визначимо поняття узагальненого образу

Х,У  ΩR: imR(X,Y)={y | zR & z[X]=x & z[Y]=y}

Визначимо Z, як Z = ΩR \(XY).

ХУ (У **багатозначно залежить** від Х – списка атрибутів), якщо

 (x,z)  XZ, imR(XZ, Y) = imR(X, Y).

Розглянемо ще одну таблицю ВС. Між її атрибутами знову таки немає функціональної залежності.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ВС | КЛ | КП | Рік | З/п |
|  | ПТ | геометрія | 1979 | 180 |
|  | ПТ | алгебра | 1979 | 180 |
|  | ПТ | геометрія | 1980 | 200 |
|  | ПТ | алгебра | 1980 | 200 |
|  | СД | математика1 | 1979 | 250 |
|  | СД | математика2 | 1979 | 250 |
|  | СД | математика1 | 1980 | 270 |
|  | СД | математика2 | 1980 | 270 |

Проаналізуємо цю реляцію щодо наявності багатозначних залежностей. Нехай Х – це КЛ, а У – КП, тоді Z – це (рік, з/п). Узагальнений образ {ПТ, 1979, 180} є {геометрія, алгебра}, узагальнений образ {ПТ} - це теж { геометрія, алгебра}. Крім того, узагальнений образ {ПТ, 1980, 200} теж є {геометрія, алгебра}. Аналогічний результат отримаємо також і для значення «СД». Таким чином, наповнення таблиця не заперечує, що КЛКП. Подібним же способом ми можемо проаналізувати наповнення і для КЛ (рік, з/п). Відзначимо, що на основі даних в таблиці можемо зробити висновок про те, що наповнення таблиці заперечує чи не заперечує наявність тої чи іншої багатозначної залежності, але не можемо стверджувати, що така залежність є. Наявність багатозначної залежності (наприклад, КЛКП) випливає з аналізу предметної області, наприклад, відомо що КП відповідає чітко визначена підмножина КЛ.

Функціональна залежність є частковим випадком багатозначної залежності, тобто, якщо має місце функціональна залежність, то є і багатозначна; зворотне твердження, взагалі кажучи, невірне.

Багатозначна залежність називається **тривіальною**, якщо вона дублюється функціональною, інакше вона є **нетривіальною**.

Реляція знаходиться в **4НФ,** якщо вона знаходиться в 3НФ і не має нетривіальної багатозначної залежності, або

4НФ: R  AB  A ΩR +3 НФ (А - квазіключ)

Реляцію ВС можна декомпонувати на реляції в 4 НФ. ВС1(КЛ,КП) і ВС2(КЛ, рік, з/п), тоді отримаємо 2 таблиці по 4 рядки у кожній, але довжина рядків менша.

Теорема Fagin. (ця теорема дуже схожа на теорему Heath).

Якщо в реляції R(А,В,С) є залежність AB, то реляція може бути декомпонована на 2 реляції R(А,В) і R(А,С) без втрат даних, тобто R може бути відновлена природним з’єднанням по А. (А, В,С – це списки атрибутів).

Розглянемо тепер ще один вид залежностей, які узагальнюють багатозначні залежності. Легко бачити, що в теоремах Heath і Fagin з існування залежності *випливає* можливість декомпозиції без втрат даних, але обернене, взагалі кажучи, невірно.

Для реляції R(А,В,С) А ~» В можлива декомпозиція на 2 реляції R(А,В) і R(А,С) без втрат даних. Такі залежності будемо називати залежностями **по з’єднанню без втрат**.

Багатозначні залежності є частковим випадком залежностей по з’єднанню без втрат, тобто якщо має місце багатозначна залежність, то є і залежність по з’єднанню без втрат; зворотне твердження, взагалі кажучи, невірне.

Залежність по з’єднанню без втрат називається **тривіальною**, якщо вона дублюється багатозначною, інакше вона є **нетривіальною**.

Реляція знаходиться в **5НФ**, якщо вона знаходиться в 4НФ і не має нетривіальних залежностей по з’єднанню без втрат.

Функціональна залежність  багатозначна залежність  залежність по з’днанню без втрат.