Лекція 14

Функціональні залежності

S

S#	SNAME	STATUS	CITY
S1	Smith	20	London
S2	Jones	10	Paris
S3	Blake	30	Paris
S4	Clark	20	London
S5	Adams	30	Athens

J

J#	JNAME	CITY
J1	Sorter	Paris
J2	Display	Rome
J3	OCR	Athens
J4	Console	Athens
J5	RAID	London
J6	EDS	Oslo
J7	Tape	London

SPJ

S#	P#	J#	QTY
S1	P1	J1	200
S1	P1	J4	700
S2	P3	J1	400
S2	P3	J2	200
S2	P3	J3	200
S2	P3	J4	500
S2	P3	J5	600
S2	P3	J6	400
S2	P3	J7	800
S2	P5	J2	100
S4	P6	J3	300
S4	P6	J1	300
S5	P2	J2	200
S5	P2	J4	100
S5	P5	J5	500
S5	P5	J7	100
S5	P6	J2	200
S5	P1	J4	100
S5	P3	J4	200
S5	P4	J4	800
S5	P5	J4	400
S5	P6	J4	500

P

P#	PNAME	COLOR	WEIGHT	CITY
P1	Nut	Red	12.0	London
P2	Bolt	Green	17.0	Paris
P3	Screw	Blue	17.0	Oslo
P4	Screw	Red	14.0	London
P5	Cain	Blue	12.0	Paris
P6	Cog	Red	19.0	London

SCP

S#	CITY	P#	QTY
S1	London	P1	100
S1	London	P2	100
S2	Paris	P1	200
S2	Paris	P2	200
S3	Paris	P2	300
S4	London	P2	400
S5	London	P4	500
S6	London	P5	600

Функціональні залежності

- По суті, функціональна залежність є зв'язком типу "багато до одного" між множинами атрибутів всередині даної змінної відношення.

Приклад функціональної залежності

- Для змінної відношення поставок SP існує функціональна залежність між множинами атрибутів {S#,P#} та {QTY}.
- Це означає, що для усякого допустимого значення цієї змінної відношення справедливі наступні правила.
  - Для будь-якої заданої пари значень атрибутів S# і P# існує тільки одне відповідне їм значення атрибуту QTY.
  - Але одне і те ж відповідне їм значення атрибуту QTY можуть мати багато різних пар значень атрибутів S# і P#.

SP

S#	P#	QTY
S1	P1	100
S1	P2	100
S2	P1	200
S2	P2	200
S3	P2	300
S4	P2	400
S5	P4	500
S6	P5	600

Визначення ФЗ для відношення

- Нехай R є відношенням, а X і Y — довільні підмножини множини атрибутів відношення R. Тоді Y функціонально залежить від X, ($X \rightarrow Y$) тоді і тільки тоді, коли кожне значення множини X відношення R зв'язано точно з одним значенням множини Y відношення R.
- Інакше кажучи, якщо два кортежі відношення R співпадають по значенню X, вони співпадають і по значенню Y.

SCP

S#	CITY	P#	QTY
S1	London	P1	100
S1	London	P2	100
S2	Paris	P1	200
S2	Paris	P2	200
S3	Paris	P2	300
S4	London	P2	400
S5	London	P4	500
S6	London	P5	600

ФЗ у відношеннях

- Відношення SCP задовольняє вимогам функціональної залежності { S# } → { CITY }, оскільки всі кортежі відношення SCP з однаковими значеннями атрибуту S# мають одне і те ж значення атрибуту CITY.

SCP

S#	CITY	P#	QTY
S1	London	P1	100
S1	London	P2	100
S2	Paris	P1	200
S2	Paris	P2	200
S3	Paris	P2	300
S4	London	P2	400
S5	London	P4	500
S6	London	P5	600

ФЗ у відношеннях

- Насправді, відношення SCP задовольняє вимогам зразу декількох функціональних залежностей:
  - {S#,P#}→{QTY}
  - {S#,P#}→{CITY}
  - {S#,P#}→{CITY,QTY}
  - {S#,P#}→{S#}
  - {S#,P#}→{S#,P#,CITY,QTY}
  - {S#} →{QTY}
  - {QTY} → {S#}

Детермінант і залежна частина

- Ліву частину ФЗ називають **детермінантом**.
- Праву частину – **залежною частиною**.

ФЗ – обмеження цілісності

- Які з ФЗ виконуються для поточного значення SCP, а які – для будь-якого значення змінної відношення SCP?
  - {S#,P#}→{QTY}
  - {S#,P#}→{CITY}
  - {S#,P#}→{CITY,QTY}
  - {S#,P#}→{S#}
  - {S#,P#}→{S#,P#,CITY,QTY}
  - {S#} →{QTY}
  - {QTY} → {S#}

SCP

S#	CITY	P#	QTY
S1	London	P1	100
S1	London	P2	100
S2	Paris	P1	200
S2	Paris	P2	200
S3	Paris	P2	300
S4	London	P2	400
S5	London	P4	500
S6	London	P5	600

- ФЗ, що виконуються завжди, є обмеженням цілісності даних для змінної відношення, оскільки дана ФЗ накладає певні обмеження на всі допустимі значення цієї змінної відношення.

Визначення ФЗ для змінної відношення

- Нехай R – **змінна** відношення, а X і Y - довільні підмножини множини атрибутів **змінної** відношення R. Тоді Y функціонально залежить від X, (X → Y) тоді і тільки тоді, коли для будь-якого допустимого значення **змінної** відношення R кожне значення множини X **змінної** відношення R зв'язано точно з одним значенням множини Y **змінної** відношення R.
- Інакше кажучи, для будь-якого допустимого значення змінної відношення R, якщо два кортежі змінної відношення R співпадають по значенню X, вони також співпадають і по значенню Y.

ФЗ і потенційні ключі

- Якщо X є потенційним ключем змінної відношення R, то всі атрибути Y змінної відношення R повинні обов'язково бути функціонально залежними від X. Це безпосередньо випливає із визначення потенційного ключа.
- Якщо ж змінна відношення R задовольняє ФЗ A→B і A не є потенційним ключем, то R обов'язково буде характеризуватися деякою збитковістю.
  - Наприклад, у змінній відношенні SCP, присутність ФЗ S# → CITY призводить до того, що дані про місце знаходження постачальника в певному місті повторюється багато разів.

Множина ФЗ

- Функціональні залежності є обмеженнями цілісності, тому бажано, щоб СКБД забезпечувала їх виконання.
- Для кожної заданої множини функціональних залежностей S бажано знайти таку множину T, яка (в ідеальній ситуації) була б суттєво меншою множини S і при цьому кожна функціональна залежність в множині S могла б бути замінена функціональною залежністю з множини T.
- Якщо б така множина T була знайдена, то СКБД достатньо було б контролювати виконання функціональних залежностей з множини T, що автоматично забезпечувало б виконання всіх функціональних залежностей з множини S.

ТРИВІАЛЬНІ ТА НЕТРИВІАЛЬНІ ЗАЛЕЖНОСТІ

- Залежність називається тривіальною, якщо вона не може не виконуватися.
- Функціональна залежність є тривіальною тоді і тільки тоді, коли права частина її символічного запису є підмножиною лівої частини.
- Кажуть, що ФЗ виду $\{A_1, A_2, \ldots, A_n\} \rightarrow \{B_1, B_2, \ldots, B_m\}$ відноситься до категорії:
  - тривіальних, якщо множина $\{B_1, B_2, \ldots, B_m\}$ є підмножиною множини $\{A_1, A_2, \ldots, A_n\}$;
  - нетривіальних, якщо принаймні один з атрибутів B_i не є елементом $\{A_1, A_2, \ldots, A_n\}$;
  - повністю нетривіальних, якщо ні один з атрибутів B_i не є елементом $\{A_1, A_2, \ldots, A_n\}$;

ЗАМИКАННЯ МНОЖИНИ ЗАЛЕЖНОСТЕЙ

- Множина всіх функціональних залежностей, які випливають з заданої множини функціональних залежностей S, називається замиканням множини S і позначається символом S+.

Аксіоми Армстронга

- Нехай $\{A_1, A_2, \ldots, A_n\} \equiv \mathbf{A}$, $\{B_1, B_2, \ldots, B_m\} \equiv \mathbf{B}$, $\{C_1, C_2, \ldots, C_m\} \equiv \mathbf{C}$ і $\{D_1, D_2, \ldots, D_m\} \equiv \mathbf{D}$ тоді:
  - Рефлексивність. Якщо $B \in A$, то $A \rightarrow B$.
  - Доповнення. Якщо $A \rightarrow B$, то $AC \rightarrow BC$.
  - Транзитивність. Якщо $A \rightarrow B$ і $B \rightarrow C$, то $A \rightarrow C$.
  - Самовизначення. $A \rightarrow A$.
  - Декомпозиція. Якщо $A \rightarrow BC$, то $A \rightarrow B$ і $A \rightarrow C$.
  - Об'єднання. Якщо $A \rightarrow B$ і $A \rightarrow C$, то $A \rightarrow BC$.
  - Композиція. Якщо $A \rightarrow B$ і $C \rightarrow D$, то $AC \rightarrow BD$.
  - Якщо $A \rightarrow B$ і $C \rightarrow D$, то $A \cup (C - B) \rightarrow BD$

Приклад використання аксіом Армстронга

- Нехай дано деяку змінну відношення R з атрибутами A, B, C, D, E, F і наступними ФЗ:
  - $A \rightarrow BC$
  - $B \rightarrow E$
  - $CD \rightarrow EF$
- Показати, що для змінної відношення R виконується також функціональна залежність $AD \rightarrow F$, яка внаслідок цього належить замиканню заданої множини функціональних залежностей.

ЗАМИКАННЯ МНОЖИНИ АТРИБУТІВ

- Замиканням множини $\{A_1, A_2, \ldots, A_n\}$ при умові виконання ФЗ множини S називається множина B атрибутів, така що для кожного відношення, якому задовольняють всі ФЗ множини S, справедлива і ФЗ $\{A_1, A_2, \ldots, A_n\} \rightarrow B$.
- Замикання множини атрибутів $\{A_1, A_2, \ldots, A_n\}$ позначається як $\{A_1, A_2, \ldots, A_n\}+$

Алгоритм пошуку замикання множини атрибутів

- Присвоїть $Z = \{A_1, A_2, \ldots, A_n\}$
- Знайти ФЗ $\{B_1, B_2, \ldots, B_k\} \rightarrow C$, таку що $\{B_1, B_2, \ldots, B_k\}$ всі належать Z, але C – не належить
- Якщо вказана ФЗ існує, C додати до Z
- Повторити, поки жоден атрибут не може бути доданий до Z
- $Z = \{A_1, A_2, \ldots, A_n\}^+$

Приклад побудови замикання атрибутів

- Дано R {A, B, C, D, E, F} та наступні ФЗ.

  $A \rightarrow BC$

  $E \rightarrow CF$

  $B \rightarrow E$

  $CD \rightarrow EF$

- Побудувати замикання {A, B}+ множини атрибутів {A, B}.

Приклад побудови замикання атрибутів

- Дано R {A, B, C, D, E, F} та наступні ФЗ.

  AB → C

  BC → AD

  D → E

  CF → B

- Побудувати замикання {A, B}+ множини атрибутів {A, B}.

Приклад використання замикання атрибутів

- Дано R {A, B, C, D, E, F} та наступні ФЗ.

  $AB \rightarrow C$

  $BC \rightarrow AD$

  $D \rightarrow E$

  $CF \rightarrow B$

- Перевірити, чи випливає з множини залежностей ФЗ $AB \rightarrow D$?
- Перевірити, чи випливає з множини залежностей ФЗ $D \rightarrow A$?

Висновки

- Для заданої множини ФЗ S легко можна вказати, чи буде задана ФЗ $X \to Y$ випливати з S, оскільки це можливо тоді і тільки тоді, коли множина Y є підмножиною замикання X+ множини X для заданої множини S.

Висновки

- Суперключ змінної відношення R – це множина атрибутів змінної відношення R, яка у вигляді подмножини містить принаймні один потенційний ключ.
- Суперключі для даної змінної відношення R – це такі підмножини K множини атрибутів змінної відношення R, що ФЗ $K \rightarrow A$ виконується для кожного атрибуту A змінної відношення R.
- Множина K є суперключем тоді і тільки тоді, коли замикання K+ для множини K в межах заданої множини ФЗ є множиною абсолютно всіх атрибутів змінної відношення R.
- $\{A_1,A_2,\ldots,A_n\}^+$ буде множиною всіх атрибутів відношення, якщо і тільки якщо $A_1,A_2,\ldots,A_n$ утворює суперключ цього відношення.

Завдання

- Для змінної відношення R{A,B,C,D,E,F,G} визначено ФЗ.

  $A \rightarrow B$

  $BC \rightarrow DE$

  $AEF \rightarrow G$
- Побудувати замикання {A,C}+ для даної множини ФЗ.
- Чи випливає з цієї множини ФЗ $ACF \rightarrow DG$?

Проекціювання ФЗ

- Нехай R – змінна відношення, для якої виконується множина ФЗ S.
- Нехай P – відношення, отримане з R за допомогою оператора проекції.
- Які ФЗ залишаться справедливими для P?

Проекціювання ФЗ

- Нехай R{A, B, C, D} – змінна відношення, має ФЗ $A \to B$, $B \to C$, $C \to D$. Необхідно побудувати S=R{A,C,D} (тут {} – реляційна проекція).
- Щоб знайти ФЗ, що задовольняють S, треба побудувати замикання для всіх восьми підмножин {A, C, D}, використавши повну множину ФЗ.
- Для кожного замикання деякої множини X можна додати ФЗ виду $X \to E$, де E – кожний атрибут, присутній в X+ і у відношенні S, але відсутній в X.

Замкнені множини ФЗ

- На основі заданої множини ФЗ зазвичай можна вивести інші ФЗ.
- Інколи існує можливість вибору одного з альтернативних наборів ФЗ, придатних для представлення повної множини ФЗ конкретного відношення.
- Будь-яка множина ФЗ змінної відношення, з якої допустимо вивести всі інші ФЗ, називається **базисом**.
- В свою чергу, якщо жодна з підмножин базису не може бути використана для отримання повної множини ФЗ, говорять, що базис є **мінімальним**.

Покриття множин ФЗ

- Нехай S_1 і S_2 — дві множини ФЗ. Якщо будь-яка ФЗ, яка випливає з множини залежностей S_1, випливає також із множини залежностей S_2 то множина S_2 називається покриттям для множини S_1.
  - Це означає, що якщо СКБД забезпечить виконання обмежень, представлених залежностями множини S_2, то автоматично будуть виконані і всі обмеження, встановлені залежностями множини S_1.

Покриття множин ФЗ

- Якщо множина S_2 є покриттям для множини S_1, а множина S_1 одночасно є покриттям для множини S_2 (тобто, якщо $S_1+=S_2+$), то множини S_1 і S_2 еквівалентні.

Мінімальний базис

- Множина функціональних залежностей називається мінімальним базисом тоді і тільки тоді, коли вона має всі три властивості:
  - Права (залежна) частина кожної ФЗ із множини S містить тільки один атрибут.
  - Ліва частина (детермінант) кожної ФЗ із множини S, в свою чергу, є мінімальною, тобто жоден атрибут із детермінанта не може бути опущений без зміни замикання S+
  - Ні одна ФЗ із множини S не може буть видалена із множини S без зміни його замикання S+.

Приклад побудови мінімального базису

- Дано змінну відношення R {A,B,C,D} з наступними ФЗ.

  $A \rightarrow BC$

  $B \rightarrow C$

  $A \rightarrow B$

  $AB \rightarrow C$

  $AC \rightarrow D$
- Побудувати мінімальний базис, еквівалентний даній множині.

Завдання

- Визначити, чи еквівалентні дві множини ФЗ, встановлених для змінної відношення R{A, B, C, D, E}.
  - $A \to B$, $AB \to C$, $D \to AC$, $D \to E$
  - $A \to BC$, $D \to AE$

Завдання

- Найти мінімальне покриття (базис) множини функціональних залежностей, заданих для змінної відношення R{ A, B, C, D, E, F}.

  $AB \to C$

  $C \to A$

  $BC \to D$

  $ACD \to B$

  $BE \to C$

  $CE \to FA$

  $CF \to BD$

  $D \to EF$

Завдання

- Нехай задана змінна відношення NADDR з атрибутами NAME (Унікальне ім’я), STREET (Вулиця), CITY (Місто), STATE (Штат) і ZIP (Поштовий індекс).
- Вважаємо, по-перше, що кожному поштовому індексу відповідає тільки одне місто і штат, по-друге, що кожній вулиці, місту і штату відповідає тільки один поштовий індекс.
- Знайти мінімальну множину ФЗ для цієї змінної відношення. Які потенційні ключі існують для цієї змінної відношення?

Завдання

- Дано R{A, B, C, D, E, F, G, H, I, J}, для якої виконується множина ФЗ.

  $ABD \to E$

  $AB \to G$

  $B \to F$

  $C \to J$

  $CJ \to I$

  $G \to H$

- Чи є ця множина мінімальною?
- Які потенційні ключі існують для даної змінної відношення?