

## 24. Бази от данни: Нормални форми

1. Проектиране схемите на релационните бази от данни  
Да намерим/....

2. Аномалии, ограничения, ключове

**Аномалиите** представляват недостатъците на конкретна релационна схема. Видовете аномалии са:

- *Аномалии от излишество на данните* - когато едни и същи данни се записват на различни места
- *Аномалии при добавяне на данните*
- *Аномалии при промяна на данните*
- *Аномалии при изтриване на данните*

Пример: Дадена е релационна схема БИБЛИОТЕКА (Биб#, Адрес, Книга, Брой), в която:

- Биб# - номер на библиотеката
- Адрес - адрес на библиотеката
- Книга - име на книгата
- Брой - брой екземпляри на книгата

Ключовите атрибути в релацията са Биб# и Книга. Всяка библиотека еднозначно е определена от своя номер. Присъстват следните аномалии:

- a. Аномалия от излишество - за всяка книга от една и съща библиотека ще има повторение на адреса на библиотеката
- b. Аномалия при добавяне - адресът на новосъздаваща се библиотека не може да бъде въведен в БД, докато не се посочи поне една книга, която ще се съхранява в тази библиотека.
- c. Аномалия при обновяване - от (a) при промяна на адреса на библиотеката ще трябва да променим толкова редове от релацията, колкото са книгите, които се съхраняват в нея.
- d. Аномалия при изтриване - Ако поради ремонт на дадена библиотека книгите са преместени другаде, ще се наложи всички редове от релацията, отнасящи се до тази библиотека, да бъдат отстранени от релацията. Като следствие от това ще бъде "изгубен" и адресът на библиотеката, независимо че тя съществува и адресът ѝ остава същият.

Ключ:

$K = \{A_1, A_2, \dots, A_n\}$  е **ключ** за релацията  $R$ , ако:

- a. Множеството  $K$  функционално определя всички атрибути на  $R$ .
- b. За нито едно подмножество на  $K$ , (a) не е вярно.

Ако  $K$  удовлетворява (a), но не и (b), то  $K$  е суперключ.

**Ограничения** дефинират правила за стойностите, допустими за съответните колони и представляват стандартен механизъм за налагане на цялостност

- a. Ограничения за ключа
  - i. Уникални стойности (UNIQUE и PRIMARY KEY)- не се допуска въвеждането на стойност, която вече съществува в колоната
  - ii. Липса на null стойности (NOT NULL) - не се допуска въвеждането на null като стойност в колоната
- b. Ограничение за референциален интегритет - стойност, сочена от някакви елементи действително съществува в базата данни
  - i. Ограничение за вторичен ключ (FOREIGN KEY)- за всяка ненулева стойност на атрибут, който е външен ключ, трябва да съществува ключов атрибут от друга релация, който съдържа тази стойност
- c. Ограничения на домейна (CHECK)
  - i. Ограничават стойностите на даден атрибут да принадлежат в дадено множество
- d. Общи ограничения - те са част от схемата

### 3. Функционални зависимости

Функционалните зависимости представляват зависимости между атрибутите в рамките на една релация.

Дефиниция на функционална зависимост:  $A_1, A_2, \dots, A_n$  функционално определят  $B$ , пишем  $A_1, A_2, \dots, A_n \rightarrow B$ :

- Ако два кортежа от екземпляр на релацията  $R$  съвпадат по атрибутите  $A_1, A_2, \dots, A_n$  от  $R$ , то те трябва да съвпадат и по атрибута  $B$ .

Нарича се функционална, защото има функция, която на списък от стойности (по една за всяко  $A_1, A_2, \dots, A_n$ ) съпоставя уникална стойност от  $B$ .

Най-основният пример за функционална зависимост е тази между ключа на релацията и всички атрибути в релацията.

### 4. Аксиоми на Армстронг

За примери ще използваме следната релация:

Id#	Name	Address	C#	Description	Grade
124	Jones	Phila	Phil7	Plato	A
456	Smith	NYC	Phil7	Plato	B
789	Brown	Boston	Math8	Topology	C
124	Jones	Phila	Math8	Topology	A
789	Brown	Boston	Eng12	Chaucer	B

1. **Рефлексивност:** Ако  $Y \subseteq X$ , то  $X \rightarrow Y$   
Пример: Name, Address  $\rightarrow$  Address
2. **Разширение**, допълнение: Ако  $X \rightarrow Y$ , то  $XW \rightarrow YW$   
Пример: От  $C\# \rightarrow Description$ , получаваме  $C\#, ID\# \rightarrow Description, ID\#$
3. **Транзитивност:** Ако  $X \rightarrow Y$  и  $Y \rightarrow Z$ , то  $X \rightarrow Z$   
Пример: от  $ID\#, C\# \rightarrow C\#$  и  $C\# \rightarrow Description$  получаваме  $ID\#, C\# \rightarrow Description$

Чрез тях могат да се докажат останалите зависимости, използваме аксиомите на Армстронг, за да създаваме нови зависимости.

Следствия от Аксиомите на Армстронг

1. **Обединение:** Ако  $X \rightarrow Y$  и  $X \rightarrow Z$ , то  $X \rightarrow YZ$
2. **Псевдотранзитивност:** Ако  $X \rightarrow Y$  и  $WY \rightarrow Z$ , то  $XW \rightarrow Z$
3. **Декомпозиция:** Ако  $X \rightarrow Y$  и  $Z \subseteq Y$ , то  $X \rightarrow Z$

Функционална зависимост  $A_1A_2 \dots A_n \rightarrow B_1B_2 \dots B_m$  се нарича

- **Тривиална**, ако атрибутите  $B_1B_2 \dots B_m$  са подмножество на  $A_1A_2 \dots A_n$   
Пример:  $title, year \rightarrow year$
- **Нетривиална:** поне един от атрибут от  $B_1B_2 \dots B_m$  не е подмножество на  $A_1A_2 \dots A_n$   
Пример:  $title, year \rightarrow year, length$
- **Напълно нетривиална:** нито един от атрибутите  $B_1B_2 \dots B_m$  не е част от  $A_1A_2 \dots A_n$

### 5. Първа, втора, трета нормална форма, нормална форма на Бойс-Код

Дефиниция на Първа нормална форма (1НФ):

- Всеки компонент на всеки кортеж съдържа атомарно значение  
Имаме ли таблица, която по нашите дефиниции е релация, то тя е в 1НФ

Дефиниция на Втора нормална форма (2НФ):

- Релация  $R$  е във 2НФ, когато е изпълнена 1НФ и всеки неключов атрибут е в пълна функционална зависимост от ключа, т.е. зависи от целия ключ, а не от някакво подмножество на ключа

Дефиниция на Трета нормална форма (3НФ)

- Атрибут, който е част от ключа се нарича **първичен атрибут**
- Релация  $R$  е в 3НФ, т.с.т.к. за всяка функционална зависимост е изпълнено, че или лявата страна е суперключ или дясната е първичен атрибут
- Релацията  $R$  е в 3НФ т.с.т.к. за всяка нетривиална зависимост  $A_1A_2 \dots A_n \rightarrow B$  от  $R$ , съответното м-во от атрибути  $\{A_1, A_2, \dots, A_n\}$  е суперключ за  $R$  или  $B$  е част от ключ.

Дефиниция на Нормална форма на Бойс-Код (BCNF)

- Релацията  $R$  е в нормална форма на Бойс-Код т.с.т.к. за всяка нетривиална зависимост  $A_1A_2 \dots A_n \rightarrow B_1B_2 \dots B_m$  от  $R$ , съответното  $m$ -во от атрибути  $\{A_1, A_2, \dots, A_n\}$  е суперключ за  $R$ .

## 6. Многозначни зависимости (MVD)

Има ситуации, при които независимостта на атрибутите не се изразява чрез функционални зависимости. Например, когато всички атрибути са част от ключа. Релацията е в BCNF, защото няма нетривиални зависимости, но все пак има излишество на информация.

$A_1A_2 \dots A_n \twoheadrightarrow B_1B_2 \dots B_m$  е **многозначна зависимост** в  $R$ , ако за всяка двойка кортежи  $t, u$  от  $R$ , за които  $t[A_1A_2 \dots A_n] = u[A_1A_2 \dots A_n]$  съществува кортеж  $v$  от  $R$ , за който:

- $v[A_1A_2 \dots A_n] = t[A_1A_2 \dots A_n] = u[A_1A_2 \dots A_n]$
- $v[B_1B_2 \dots B_m] = t[B_1B_2 \dots B_m]$
- $v[C_1C_2 \dots C_k] = u[C_1C_2 \dots C_k]$ ,

където  $C_1C_2 \dots C_k$  са всички атрибути в  $R$  с изключение на  $(A_1A_2 \dots A_n \cup B_1B_2 \dots B_m)$ .

$X \twoheadrightarrow Y$  утвърждава, че ако 2 кортежа в една релация съвпадат по всички атрибути на  $X$ , техните компоненти от множеството атрибути  $Y$  могат да бъдат разменени и резултатът ще даде 2 нови кортежа, които също принадлежат на релацията

## 7. Аксиоми на многозначните зависимости

**Тривиална MVD:**

- $A_1A_2 \dots A_n \rightarrow B_1B_2 \dots B_m$ , когато  $B_1B_2 \dots B_m$  е подмножество на  $A_1A_2 \dots A_n$  или  $(A_1A_2 \dots A_n \cup B_1B_2 \dots B_m)$  съдържа всички атрибути на  $R$

**Нетривиална MVD:**

- Когато нито еидн от атрибутите  $B_1B_2 \dots B_m$  не съвпада с  $A_1A_2 \dots A_n$
- Не всички атрибути на  $R$  принадлежат на  $(A_1A_2 \dots A_n \cup B_1B_2 \dots B_m)$

**Транзитивно правило:**

- Ако  $A_1A_2 \dots A_n \twoheadrightarrow B_1B_2 \dots B_m$  и  $B_1B_2 \dots B_m \twoheadrightarrow C_1C_2 \dots C_k$ , то  $A_1A_2 \dots A_n \twoheadrightarrow C_1C_2 \dots C_k$

**Правило на допълнението:**

- Ако  $A_1A_2 \dots A_n \twoheadrightarrow B_1B_2 \dots B_m$ , то  $A_1A_2 \dots A_n \twoheadrightarrow C_1C_2 \dots C_k$ , където  $C_1C_2 \dots C_k$  е множеството от всички атрибути на  $R$  с изключение на  $(A_1A_2 \dots A_n \cup B_1B_2 \dots B_m)$ .

**Правило на обединението**

- Ако  $X_1X_2 \dots X_n \twoheadrightarrow Y_1Y_2 \dots Y_m$  и  $X_1X_2 \dots X_n \twoheadrightarrow Z_1Z_2 \dots Z_k$ , то  $X_1X_2 \dots X_n \twoheadrightarrow (Y_1Y_2 \dots Y_m \cup Z_1Z_2 \dots Z_k)$

**Правило FD-IS-AN-MVD (всяка функционална зависимост е многозначна зависимост):**

- Ако  $A_1A_2 \dots A_n \rightarrow B_1B_2 \dots B_m$ , то  $A_1A_2 \dots A_n \twoheadrightarrow B_1B_2 \dots B_m$ .
- Доказателство:

Да предположим, че  $R$  е релация, за която е вярно  $A_1A_2 \dots A_n \rightarrow B_1B_2 \dots B_m$  и да предположим, че  $t, u$  са кортежи, които съвпадат в  $A$ . За да покажем, че е вярно  $A_1A_2 \dots A_n \twoheadrightarrow B_1B_2 \dots B_m$  трябва да докажем, че  $R$  също съдържа кортеж  $v$ , който са изпълнени условията от дефиницията на MVD. Но кортежът  $v$  може да е  $u$ . Знаем, че  $u$  съвпада с  $t$  и  $u$  в  $A$ . ФЗ  $A_1A_2 \dots A_n \rightarrow B_1B_2 \dots B_m$  предполага, че  $u$  съвпада с  $t$  в  $B$ , и се съгласува със себе си по останалите атрибути, така получихме, че когато е в сила функционалната зависимост, то е в сила и многозначната зависимост.