

Многозначни зависимости 4 нормална форма

ОСНОВНИ ТЕМИ

- Дефиниция на многозначни зависимости (Multivalued Dependencies – MDs)
- Причини за появата на MDs
- 4 нормална форма (4NF)

Многозначни зависимости

- 2 атрибута или множество атрибути са независими помежду си
- Обобщение на FDs
- Наличие на ситуации, при които независимостта на атрибутите не се изразява чрез FDs

Мотивация

- Съществуват схеми в BCNF, които съдържат излишни данни

Stars

<i>name</i>	<i>street</i>	<i>city</i>	<i>title</i>	<i>year</i>
C. Fisher	123 Maple Str.	Hollywood	Star Wars	1977
C. Fisher	5 Locust Ln.	Malibu	Star Wars	1977
C. Fisher	123 Maple Str.	Hollywood	Empire Strikes Back	1980
C. Fisher	5 Locust Ln.	Malibu	Empire Strikes Back	1980
C. Fisher	123 Maple Str.	Hollywood	Return of the Jedi	1983
C. Fisher	5 Locust Ln.	Malibu	Return of the Jedi	1983

Независимост на атрибути

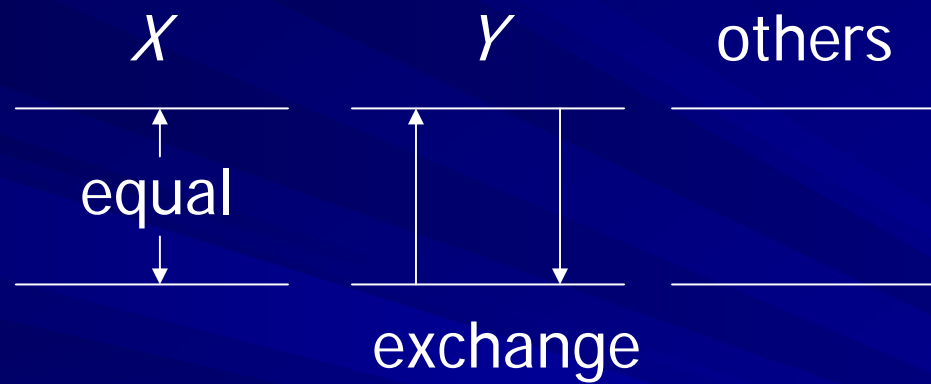
- Няма основание да свържем адреса с 1 филм и да не го свържем с друг
- При повторенията на адресите и фактите за филмите във всички техни комбинации получаваме излишество на данни
- Няма нарушение на BCNF
 - Не съществуват не-тривиални FD's- 5-те атрибута формират 1! суперключ
 - Защо?

Дефиниция на MVD

■ *Multivalued dependency* (MVD)

$X \twoheadrightarrow Y$ 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

MVD $X \rightarrow Y$



Multivalued dependency - дефиниция

Дефиниция: Многозначна зависимост (MVD):

$A_1A_2...A_n \twoheadrightarrow B_1B_2...B_m$ е многозначна зависимост в R ако:

за всяка двойка кортежи t, u от R , за които

$$t[A_1A_2...A_n] = u[A_1A_2...A_n],$$

съществува кортеж v от R за който:

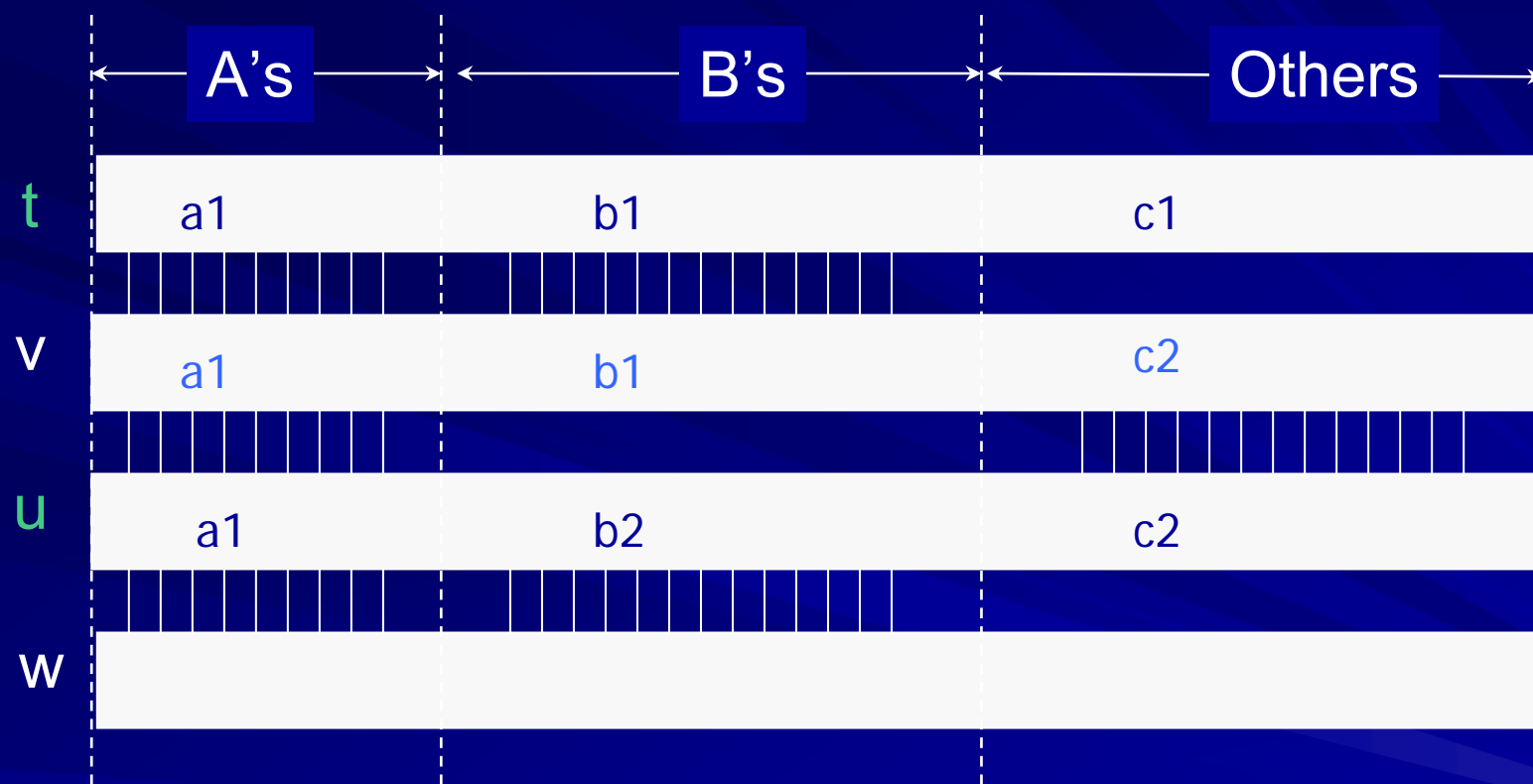
$$(1) v[A_1A_2...A_n] = t[A_1A_2...A_n] = u[A_1A_2...A_n]$$

$$(2) v[B_1B_2...B_m] = t[B_1B_2...B_m]$$

$$(3) v[C_1C_2...C_k] = u[C_1C_2...C_k],$$

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

Графично представяне



- MVD гарантира съществуването на кортежа v exists
 - Съществуването на 4 кортеж w – размяна на t и u

Пример: name →→ street city

Stars

	<i>name</i>	<i>street</i>	<i>city</i>	<i>title</i>	<i>year</i>
t	C. Fisher	123 Maple Str.	Hollywood	Star Wars	1977
	C. Fisher	5 Locust Ln.	Malibu	Star Wars	1977
v	C. Fisher	123 Maple Str.	Hollywood	Empire Strikes Back	1980
u	C. Fisher	5 Locust Ln.	Malibu	Empire Strikes Back	1980
	C. Fisher	123 Maple Str.	Hollywood	Return of the Jedi	1983
	C. Fisher	5 Locust Ln.	Malibu	Return of the Jedi	1983

Използване на MVDs

- Тривиални зависимости
- Правило за транзитивност
- Правило за попълнение
- Правило за обединение

Тривиални MVDs

■ Тривиална MVD

$$A_1A_2\dots A_n \twoheadrightarrow B_1B_2\dots B_m$$

- 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

$$A_1A_2\dots A_n \twoheadrightarrow B_1B_2\dots B_m$$

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

Правила за MVDs

■ Транзитивно правило

Ако $A_1A_2\dots A_n \twoheadrightarrow V_1V_2\dots V_m$ и $V_1V_2\dots V_m \twoheadrightarrow C_1C_2\dots C_k$, то $A_1A_2\dots A_n \twoheadrightarrow C_1C_2\dots C_k$

■ Правило на допълнението

Ако $A_1A_2\dots A_n \twoheadrightarrow V_1V_2\dots V_m$, то $A_1A_2\dots A_n \twoheadrightarrow C_1C_2\dots C_k$, където $C_1C_2\dots C_k$ е м-то от всички атрибути на R с изключение на $(A_1A_2\dots A_n \cup V_1V_2\dots V_m)$

■ 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 $X_1X_2\dots X_n \twoheadrightarrow (Y_1Y_2\dots Y_m \cup Z_1Z_2\dots Z_k)$

Свойства на MVDs

- Подобно на FD's, не можем да разделяме лявата част на MVD.
- За разлика от FD's, не можем да разделяме и дясната част --- понякога се налага да оставяме няколко атрибута в дясната част

Пример

- Stars: $\text{name} \rightarrow \rightarrow \text{street city}$
- $\text{name} \rightarrow \rightarrow \text{street ?}$

<i>name</i>	<i>street</i>	<i>city</i>	<i>title</i>	<i>year</i>
C. Fisher	5 Locust Ln.	Hollywood	Star Wars	1977
C. Fisher	5 Locust Ln.	Malibu	Star Wars	1977

Всяка FDs е MVDs

- FD-IS-AN-MVD правило

Ако $A_1A_2\dots A_n \rightarrow B_1B_2\dots B_m$ то

$$A_1A_2\dots A_n \twoheadrightarrow B_1B_2\dots B_m$$

- Докажете чрез (1), (2), (3) от MVD дефиницията

Правило на допълнението

Stars:

name →→ street city

- Съгласно правилото на допълнението

name →→ title year

Четвърта нормална форма – 4NF

- Излишеството на данни, което произтича от MVD's не може да се отстрани чрез привеждане в BCNF.
- Необходима е по-строга нормална форма, наречена 4NF, която третира MVD's като FD's по отношение на декомпозицията, но не и по отношение на ключовете.

4NF- дефиниция

- Релацията R удовлетворява 4NF, ако за всяка нетривиална MVD

$$A_1A_2\dots A_n \twoheadrightarrow V_1V_2\dots V_m ,$$

$A_1A_2\dots A_n$ е суперключ.

– Понятието ключ се основава на дефиницията на FD

- 4NF е обобщение на BCNF

– тъй като FD е MVD, $4NF \rightarrow BCNF$

BCNF и 4NF

- Всяка FD $X \rightarrow Y$ е също така и MVD, $X \twoheadrightarrow Y$.
- Ако R удовлетворява 4NF, тя очевидно удовлетворява и BCNF.
 - Всяко нарушение на BCNF е нарушение и на 4NF.
- Но R може да бъде в BCNF и да не бъде в 4NF, защото MVD's са “неуловими” от BCNF.

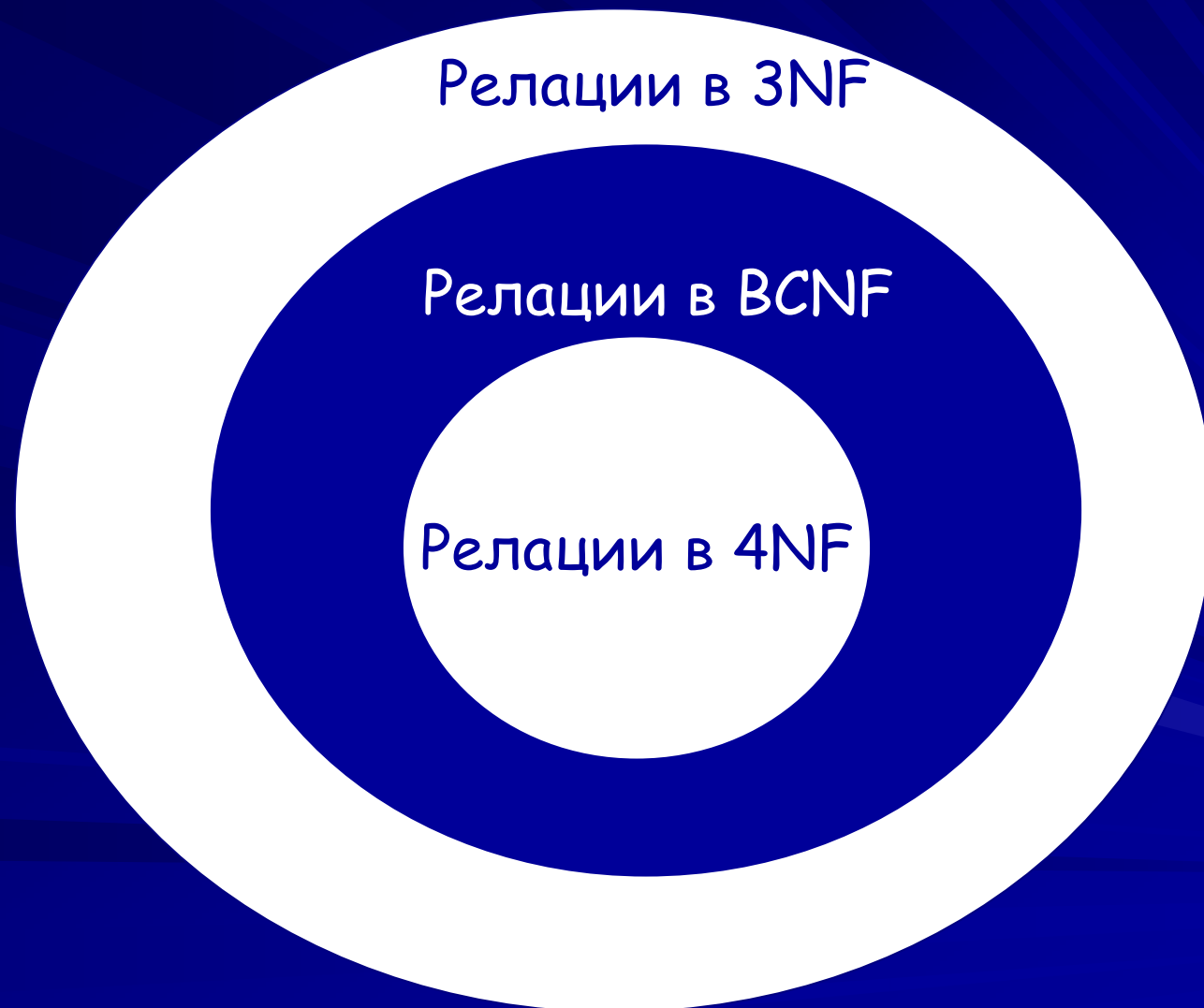
Декомпозиция и 4NF

- Ако $X \twoheadrightarrow Y$ нарушава 4NF за релацията R , извършваме декомпозиция на R , използвайки същата техника както при BCNF.
 1. XY е едната от декомпозираните релации.
 2. Всички атрибути, без $Y \cup X$ - другата.

Пример: Stars не е в 4NF

- $\text{name} \twoheadrightarrow \text{street city}$
 - Нетривиална MVD
 - Name не е суперключ
- Прилагаме декомпозиция:
R(name, street, city)
S(name, title, year)
- $\text{name} \twoheadrightarrow \text{street city}$ in R ?
- $\text{name} \twoheadrightarrow \text{title year}$ in S?

Взаимоотношения между нормалните форми



Свойства на нормалните форми

Свойство	3НФ	BCNF	4НФ
Отсъствие на FD излишество	В повечето случаи	Да	Да
Отсъствие на MVD излишество	Не	Не	Да
Запазване на FD	Да	Не винаги	Не винаги
Запазване на MVD	Не винаги	Не винаги	Не винаги