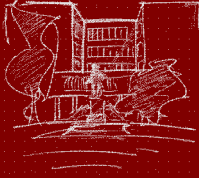


[P273] **Пројектовање база података** 7



**Саша Малков**  
Универзитет у Београду  
Математички факултет  
2023/2024


[P273] **Пројектовање база података**  
Саша Малков



Тема 8.3  
**Логичко моделирање**  
-  
**Пречишћавање схеме**

[P273] Пројектовање база података - Саша Малков - 2023/24 - час 7 1

Логичко моделирање – Пречишћавање схеме




**Пречишћавање схеме**

- Подсетимо се, логички модел се прави итеративно:
  - 1. корак: превођење логичког модела у логички
    - Прва итерација се прави на основу концептуалног модела
  - 2. корак: пречишћавање схеме, итеративно
    - Свака наредна итерација се прави мењањем претходне (енгл. *flexing*)

Универзитет у Београду – Математички факултет

[P273] Пројектовање база података - Саша Малков - 2023/24 - час 7 2

Логичко моделирање – Пречишћавање схеме



**Превођење концептуалног модела у логички**

- Видели смо како се концептуални модел преводи у логички модел (за релациони модел података)
- Примарни циљ превођења концептуалног модела у релациони:
  - Прилагођавање модела домена конкретном изабраном моделу података (нпр. релационом моделу)
  - Потпуност, тј. задржавање свих значајних односа
- Резултат превођења, међутим, често није до краја дотеран
  - Потребно је додатно пречишћавање да би био потпуно у складу са моделом података

Универзитет у Београду – Математички факултет

[P273] Пројектовање база података - Саша Малков - 2023/24 - час 7 3



## Пречишћавање схеме

- При пречишћавању схеме
  - Анализирамо добијени логички модел
  - За све појединачне елементе проверавамо усклађеност са основним концептима одговарајућег модела података
  - По потреби уводимо измене у модел
  - Водимо рачуна о захтевима (пре свега функционалним)



## Циљеви пречишћавања схеме

- Основни циљ је доследно и потпуно усклађивање модела са моделираним доменом, познатим захтевима и теоријским моделом базе података
- Критеријуми су бројни:
  - Да ли модел испуњава функционалне захтеве?
  - Да ли модел испуњава нефункционалне захтеве?
  - Да ли је модел комплетан?
  - Да ли је у складу са изабраним моделом података (нпр. релационим)?
  - Да ли модел гарантује интегритет података?
  - Да ли модел пружа потребну флексибилност?
  - Да ли модел омогућава ефикасан рад?
  - Да ли је модел довољно употребљив?
- Нису сви једнако важни
  - зависно од модела података
  - зависно од конкретног случаја



## Пречишћавање схеме и релациони модел

- У случају релационог модела при пречишћавању схеме се највише бавимо оним аспектима пројектовања који су од примарног значаја за релациони модел
  - Елиминација *редундантности*
  - Доследна примена правила интегритета
- Као основно средство користимо *функционалне зависности*



## Проблеми редундантности

- Редундантно чување података
  - неки подаци се поновљено чувају, па више копија истог податка непотребно оптерећује просторне капацитете
- Аномалије ажурирања
  - ако се једна копија поновљеног податка ажурира, база података ће бити неконзистентна ако се не ажурирају и остале копије
- Аномалије додавања
  - записивање неких података може да доводи до неконзистентности ако се не запишу још неки додатни подаци
- Аномалије брисања
  - брисање неких података може да доводи до неконзистентности ако се не обришу још неки подаци



## Пример редунадантности

- Уписан Курс ( број индекса, име и презиме студента, шифра предмета, назив предмета, име наставника )
- Постоји потенцијалан проблем са редунадантним подацима о наставницима, студентима и предметима



...

211/2015	Горан Петровић	M101	Анализа 1	Милан Марић
211/2015	Горан Петровић	M102	Линеарна алгебра	Петар Симић
212/2015	Тања Митровић	M101	Анализа 1	Милан Марић
212/2015	Тања Митровић	M102	Линеарна алгебра	Петар Симић
211/2015	Тања Митровић	M101	Линеарна алгебра	Милена Савић

- Последњи ред је неисправан
  - наведен је као илустрација могућих последица редунадантности
  - биће игнорисан у неким корацима



## Недефинисане вредности и редунадантност

- Недефинисане вредности могу да помогну у решавању неких аномалија додавања и брисања
  - Ако морамо да додамо податак а не знамо друге потребне податке, онда можемо да наведемо недефинисане вредности
    - нпр. ако студент уписује курс а не знамо наставника, можемо да наведемо да је наставник недефинисан
  - Ако хоћемо да обришемо неке редунадантне податке, а да не бришемо целе редове, онда можемо да уместо брисања упишемо недефинисане вредности
    - нпр. ако бришемо податак о наставнику на курсу а не желимо да обришемо податке о студентима који су уписали курс, можемо да упишемо да је наставник недефинисан
- ...али могу да направе неке нове проблеме...



## Декомпозиција

- Редунадантност се појављује на местима на којима постоје *не сасвим природне* везе између атрибута
- Уобичајено и *исправно* средство за решавање редунадантности је декомпозиција
- Идеја је да се једна релација са много атрибута замени већим бројем релација са по мање атрибута



...

- На пример, можемо да поделимо релацију:
  - УписанКурс( број индекса, име и презиме студента, шифра предмета, назив предмета, име наставника )
- на релације:
  - УписанКурс( број индекса, име и презиме студента, шифра предмета, назив предмета )
  - НаставникКурса( шифра предмета, име наставника )
- (тако смо решили само део проблема)



...

211/2015	Горан Петровић	M101	Анализа 1
211/2015	Горан Петровић	M102	Линеарна алгебра
212/2015	Тања Митровић	M101	Анализа 1
212/2015	Тања Митровић	M102	Линеарна алгебра

M101	Милан Марић
M102	Петар Симић



## Питања у вези декомпозиције

- Сваки пут када размишљамо о декомпозицији морамо да се запитамо:
  - Који проблем желимо да решимо предложеном декомпозицијом?
  - Да ли га заиста и решавамо?
  - Да ли тако производимо неке нове проблеме?
- Одговоре на ова питања дају
  - анализа **функционалних зависности**
  - теоријске основе **нормалних форми** релација
- ... зато ћемо прво да обрадимо одговарајуће теме...



## Теме...

- У оквиру разматрања пречишћавања схеме за релациони модел података обрадићемо следеће теме:
  - функционлне зависности
  - основне нормалне форме
  - вишезначне зависности
  - зависности спајања
  - више нормалне форме
  - поступак нормализације
- Претпоставља се да су функционалне зависности и основне нормалне форме већ изучаване у оквиру ранијих курсева...

# Пројектовање база података

[P273]

Саша Малков



Тема 8.4

## Логичко моделирање - Функционалне зависности

## Функционалне зависности, дефиниција



- **Функционална зависност** (ФЗ) је однос између два скупа атрибута на скупу торки једне релације:
- Нека је  $R$  релација и нека су  $X$  и  $Y$  непразни скупови атрибута релације  $R$ .
- Кажемо да скуп торки  $r$  релације  $R$  задовољава функционалну зависност  $X \rightarrow Y$  ако за сваки пар торки  $t1$  и  $t2$  из  $r$  важи:
  - $t1.X = t2.X \Rightarrow t1.Y = t2.Y$

## Функционалне зависности, пример



- Из дефиниције видимо да се функционалне зависности дефинишу на скуповима торки
- можемо да кажемо да представљају уопштење концепта кључа на скупу торки

## Функционалне зависности, пример



Индекс	Име	Презиме	Шиф.п.	Назив предмета	Име наставника
211/2015	Горан	Петровић	M101	Анализа 1	Милан Марић
211/2015	Горан	Петровић	M102	Линеарна алгебра	Петар Симић
212/2015	Тања	Митровић	M101	Анализа 1	Милан Марић
212/2015	Тања	Митровић	M102	Линеарна алгебра	Јована Антић
212/2015	Тања	Митровић	P100	Програмирање 1	Драгана Перић

- Које функционалне зависности можемо да уочимо на наведеном скупу торки?
  - {Индекс}  $\rightarrow$  {Име, Презиме}
  - {Шифра предмета}  $\rightarrow$  {Назив предмета}
  - {Индекс, Шифра предмета}  $\rightarrow$  {Име наставника}
  - {Име, Шифра предмета}  $\rightarrow$  {Име наставника}
  - {Презиме, Назив предмета}  $\rightarrow$  {Име наставника}
  - ... и још транзитивних зависности



## Функц. зависности на релацијама

- Ако говоримо о релацијама, онда имамо различит смисао појма зависности у вези (1) садржаја релације и (2) схеме релације:
  - на конкретним подацима *уочавамо* зависности...
  - ...т.ј. по дефиницији на скупу свих торки релације
  - на схеми релације *прописујемо* да је садржај релације *исправан* ако га чини скуп редова за који важе дате функционалне зависности
- Функционална зависност на нивоу релације базе података представља *правило иншиеришеша*
  - установљена ФЗ *мора* увек да важи на скупу свих редова релације



## Функционалне зависности, пример

Индекс	Име	Презиме	Шиф.п	Назив предмета	Име наставника
211/2015	Горан	Петровић	M101	Анализа 1	Милан Марић
211/2015	Горан	Петровић	M102	Линеарна алгебра	Петар Симић
212/2015	Тања	Митровић	M101	Анализа 1	Милан Марић
212/2015	Тања	Митровић	M102	Линеарна алгебра	Јована Антић
212/2015	Тања	Митровић	P100	Програмирање 1	Драгана Перић

- Неке ФЗ важе на посматраном скупу редова, али не представљају опште правило
- Од претходно уочених ФЗ на посматраном скупу торки неке не могу да постану ФЗ на нивоу релације:
  - **{Име, Шифра предмета} → {Име наставника}**
    - два студента са истим именом могу да слушају предмет код различитих наставника
  - **{Презиме, Назив предмета} → {Име наставника}**
    - два студента са истим презименом могу да слушају предмет код различитих наставника
    - два предмета са истим називом могу имати различите наставнике на различитим групама
  - ... и друге



## Функц. зависности на релацијама (2)

- ФЗ на релацијама установљавамо
  - разматрањем ФЗ уочених на примерима скупова података
  - разматрањем општих пословних правила која важе у домену
  - свођењем на основни скуп ФЗ које се не могу извести из других ФЗ
- Зато нам је важно да познајемо особине ФЗ



## "Аксиоме" извођења ФЗ

- Аксиоме извођења ФЗ (називају се и Армстронгове аксиоме) су комплетне и затворене на скупу свих ф.зависности
  - Називају се *аксиоме извођења* зато што се све сложеније ФЗ могу извести из основних ФЗ применом аксиома извођења
  - Заправо нису аксиоме, већ могу да се тривијално изведу из дефиниције ФЗ
  - Важе еквивалентне аксиоме на скупу торки



## "Аксиоме" извођења ФЗ

- За све скупове атрибута  $X$ ,  $Y$  и  $Z$  неке релације важе:

- **Рефлексивност**

$$X \rightarrow X$$

- користи се и у еквивалентном облику

$$Y \subseteq X \Rightarrow X \rightarrow Y$$

- **Проширивост**

$$X \rightarrow Y \Rightarrow (\forall Z) XZ \rightarrow Y$$

- користи се и у еквивалентном облику

$$X \rightarrow Y \Rightarrow (\forall Z) XZ \rightarrow YZ$$

- **Транзитивност**

$$X \rightarrow Y \wedge Y \rightarrow Z \Rightarrow X \rightarrow Z$$



## Функционалне зависности, пример

Индекс	Име	Презиме	Шиф.п	Назив предмета	Име наставника
211/2015	Горан	Петровић	M101	Анализа 1	Милан Марић
211/2015	Горан	Петровић	M102	Линеарна алгебра	Петар Симић
212/2015	Тања	Митровић	M101	Анализа 1	Милан Марић
212/2015	Тања	Митровић	M102	Линеарна алгебра	Јована Антић
212/2015	Тања	Митровић	P100	Програмирање 1	Драгана Перић

- Примери **рефлексивности**:

- {Индекс, Име, Презиме, Назив предмета} → {Презиме}
- {Индекс, Име, Презиме, Назив предмета} → {Индекс, Име, Презиме}
- {Индекс, Име, Презиме, Назив предмета} → {Индекс, Назив предмета}
- ... и друге



## Функционалне зависности, пример

Индекс	Име	Презиме	Шиф.п	Назив предмета	Име наставника
211/2015	Горан	Петровић	M101	Анализа 1	Милан Марић
211/2015	Горан	Петровић	M102	Линеарна алгебра	Петар Симић
212/2015	Тања	Митровић	M101	Анализа 1	Милан Марић
212/2015	Тања	Митровић	M102	Линеарна алгебра	Јована Антић
212/2015	Тања	Митровић	P100	Програмирање 1	Драгана Перић

- Примери **проширивости** (на скупу торки):

- {Име} → {Презиме} ⇒ {Име, Индекс} → {Презиме, Индекс}
- {Шиф.п.} → {Назив предмета} ⇒ {Шиф.п., Индекс} → {Назив предмета, Индекс}
- ... и друге



## Функционалне зависности, пример

Индекс	Име	Презиме	Шиф.п	Назив предмета	Име наставника
211/2015	Горан	Петровић	M101	Анализа 1	Милан Марић
211/2015	Горан	Петровић	M102	Линеарна алгебра	Петар Симић
212/2015	Тања	Митровић	M101	Анализа 1	Милан Марић
212/2015	Тања	Митровић	M102	Линеарна алгебра	Јована Антић
212/2015	Тања	Митровић	P100	Програмирање 1	Драгана Перић

- Примери **транзитивности** (на скупу торки):

- {Име, Шиф.п.} → {Презиме, Шиф.п.},  
{Презиме, Шиф.п.} → {Име наставника}  
⇒ {Име, Шиф.п.} → {Име наставника}
- ... и друге



## Најважније изведене особине ФЗ

- Изводе се из аксиома
  - Често се и оне називају аксиомама
- Најважније су:
  - *Адитивност* (или *унија*)
  - *Пројективност* (или *декомпозиција*)
  - *Псеудотранзитивност*



## Адитивност

- **Адитивност**

$$X \rightarrow Y \wedge X \rightarrow Z \Rightarrow X \rightarrow YZ$$

- Ако не би важило, онда би за неко  $X=x$  пројекција на  $YZ$  имала две торке, па би онда бар једна од пројекција на  $Y$  и  $Z$  такође имала две торке, а то је супротно претпоставци...



## Пројективност

- **Пројективност**

$$X \rightarrow YZ \Rightarrow X \rightarrow Y \wedge X \rightarrow Z$$
  - За свако  $X=x$  у  $YZ$  одговара тачно једна торка, па онда и пројекција на  $Y$  или  $Z$  такође има тачно једну торку...



## Псеудотранзитивност

- **Псеудотранзитивност**

$$X \rightarrow Y \wedge YZ \rightarrow W \Rightarrow XZ \rightarrow W$$

- $Z \rightarrow Z$ , из рефлексивности
- $XZ \rightarrow Z$ , из рефлексивности
- $XZ \rightarrow Y$ , на основу претпоставке и проширивости
- $XZ \rightarrow YZ$ , на основу претходна два корака и адитивности
- $XZ \rightarrow W$ , на основу претходног, претпоставке и транзитивности

- Често се користи у нешто основнијем облику:

$$X \rightarrow Y \wedge XY \rightarrow Z \Rightarrow X \rightarrow Z$$





## Значај функционалних зависности

- Значај ФЗ је вишеструк, али за нас је најважније да:
  - постојање сложених ФЗ може да укаже на редувантности
    - на пример, ако у релацији има транзитивних ФЗ, онда су испуњени предуслови за појаву редувантног садржаја релације
  - постојање ФЗ нам указује на могућност да се релација декомонује без губитка информација
    - на пример, раздвајањем транзитивних ФЗ у различите релације



## Функц. зависности и редувантност

- Нека су  $X, Y, Z$  подскупови атрибута релације  $R$
- Нека важе ФЗ:  $X \rightarrow Y$  и  $Y \rightarrow Z$ 
  - (приметимо да онда важи и  $X \rightarrow Z$ )
- Ако бар једна од ових ФЗ није 1-1, онда у релацији  $R$  могу да се појаве редувантне вредности
- На пример, нека  $X \rightarrow Y$  није 1-1, онда релација може да садржи торке (пројектоване на  $XYZ$ ):
  - $(x_1, y_1, z_1)$
  - $(x_2, y_1, z_1)$
  - вредности атрибута  $Z$  се редувантно понављају и морају да се мењају када год се мења  $Y$



## Функц. зависности и редувантност (2)

- У начелу, скоро свака релација у којој важи више функционалних зависности представља могућност за појављивање редувантних торки
  - осим ако су све те ФЗ 1-1
- Из угла редувантности, у идеалном случају би у свакој релацији требало да постоји тачно једна ФЗ
  - Улога релације би онда била да моделира ту једну ФЗ
  - У пракси то баш не може тако
    - Олакшава се исправност ажурирања
    - Али често мора да се ажурира више релација
    - Оптерећује из угла упита, због много спајања
    - Није увек ни теоријски оствариво
- Да ли можемо релацију са више ФЗ да поделимо на више релација а да не изгубимо информације?
  - (одговор на ово питање ћемо потражити за неколико минута)



## Затворење скупа функц. зависности

- У разматрању ФЗ често је потребно сагледати међузависности између скупова зависности или између скупова атрибута
- Нека је  $F$  неки скуп ФЗ релације  $R$ . **Затворење од  $F$** , у ознаци  $F^+$ , је најмањи скуп ФЗ који садржи  $F$ , такав да применом Армстронгових аксиома на  $F^+$  не може да се добије ниједна ФЗ која није већ у  $F^+$ .
- Нека је  $F$  неки скуп ФЗ релације  $R$  и нека је  $X$  неки скуп атрибута те релације. Неки скуп атрибута је **затворење од  $X$  над  $F$** , у ознаци  $X^+$ , ако постоји ФЗ  $X \rightarrow X^+$  у  $F^+$  и ако за сваку другу ФЗ  $X \rightarrow Z$  из  $F^+$  важи  $Z \subseteq X^+$ .
  - т.ј.  $X^+$  је максималан скуп атрибута зависних од  $X$



## Скуп зависних атрибута

- **Скуп зависних атрибута**  $X^+$  неког скупа атрибута  $X$  је скуп свих атрибута који су (посредно или непосредно) функционално зависни од атрибута из  $X$ 
  - т.ј. представља затворање у односу на све ФЗ у релацији
- (Алгоритме за конструкцију затворања ФЗ и атрибута научити из књиге ГПЛ)



## Функционалне зависности и интегритет

- Концепт интегритета у релационом моделу је тесно повезан са функционалним зависностима
- Видећемо како се кључ-кандидат, примарни кључ и јединствени кључ повезују са функционалним зависностима



## Наткључ и кључ-кандидат

- Скуп атрибута  $X$  релације  $R$  је наткључ ако је  $X^+ = Attr(R)$ 
  - т.ј. неки скуп атрибута можемо да користимо као кључ за идентификовање торки релације ако су сви остали атрибути релације (непосредно или посредно) зависни од њега
- Наткључ  $X$  релације  $R$  је кључ-кандидат ако за сваки други наткључ  $Y$  важи:  $Y \subseteq X \Rightarrow Y^+ = X^+$ .
  - т.ј. неки наткључ је кључ-кандидат ако не постоји наткључ који је његов прави подскуп
  - приметимо да може постојати и други наткључ и други кључ кандидат који нису подскупови од  $X$



## Интегритет кључа

- Услову кључа одговара функционална зависност облика:
  - { атрибути кључа }  $\rightarrow$  { сви остали атрибути }
- Приметимо да је ово проблематично ако кључ садржи атрибуте који могу да имају недефинисане вредности
  - ФЗ се дефинишу на поређењу, тј.
    - ако су једнаки кључеви, онда су једнаки и остали атрибути
  - Како се пореде недефинисане вредности?



## Интегритет јединствености ентитета

- Због тога се интегритет јединствености ентитета дефинише тако да представља додатно проширење и формализацију интегритета кључа
- За примарни кључ се бира кључ кандидат који не сме да има недефинисане вредности:
  1. { **атрибути примарног кључа** } → { **сви остали атрибути** }
  2. атрибути примарног кључа не смеју да имају недефинисане вредности



## Интегритет јединственог кључа

- Интегритет јединственог кључа је потенцијално слабији облик јединствености од интегритета примарног кључа
- Имамо три алтернативна облика:
  1. { **атрибути јединственог кључа** } → { **сви остали атрибути** }
  - 2а. атрибути јединственог кључа не смеју да имају недефинисане вредности
    - исто као и јединственост ентитета
  - 2б. ако атрибути јединственог кључа имају недефинисане вредности, сматра се да су оне међусобно идентичне
    - различити редови не могу да садрже исте вредности дефинисаних атрибута и исте комбинације недефинисаних атрибута
  - 2в. ако атрибути јединственог кључа имају недефинисане вредности, сматра се да су оне међусобно различите
    - различити редови могу да садрже исте вредности дефинисаних атрибута и исте комбинације недефинисаних атрибута
    - као да се не проверава јединственост кључева који имају недефинисане атрибуте



## Декомпозиције релационе схеме

- Декомпозиција релационе схеме је замењивање релационе схеме двама (или више) релационим схемама које заједно садрже све атрибуте полазне схеме
  - Обично се говори о релацијама, а мисли се на релационе схеме



## Потпуност декомпозиције

- Декомпозиција је **пошћуна** (комплетна) ако гарантује да ће одговарајуће релације увек садржати све оне информације које би садржала полазна релација
- Потпуност се формализује преко **реконструкцибилности**
  - Кажемо да је декомпозиција **пошћуна** ако полазна релација може да се реконструише спајањем резултујућих релација



## Очување зависности

- Потпуност декомпозиције може да се сведе на очување ФЗ
  - Декомпозиција **чува функционалне зависности** ако се из ФЗ на декомпонованим релацијама могу реконструисати све ФЗ на полазној релацији
- Теорема:  
**Ако декомпозиција чува функционалне зависности, онда је она потпуна.**
- Представља један од најважнијих теоријских елемената релационог модела и основ за пречишћавање схеме и елиминацију редундантности из модела
  - приметимо да је импликација само у једном смеру
    - декомпозиција може да буде потпуна а да не чува ФЗ
    - ... више касније...



## Функц. зависности на релацијама (2)

- Теорема: Нека је  $Attr(R)$  скуп атрибута релације  $R$  и нека за непразне поскупове атрибута  $X, Y \subseteq Attr(R)$  важи ФЗ  $X \rightarrow Y$ , тада важи једнакост  $R = R[XY] * R[XZ]$ , где је  $Z = Attr(R) \setminus XY$ .
  - т.ј. ако из релације издвојимо нову релацију тако да моделира једну ФЗ, а при томе у релацији задржимо атрибуте домена зависности, онда је декомпозиција потпуна
  - (формалан доказ је у књизи ГПЛ)
  - идеја је да на основу нове релације можемо да реконструисамо недостајуће атрибуте
    - али се не гарантује очување ФЗ



## Да ли се проблеми услед редундантности виде у концептуалном моделу?

- Неке зависности нису очигледне у концептуалном моделу
- Последица је да након превођења може да се добије логички модел у коме постоји редундантност
- На пример, можемо да добијемо релације:
  - $Студент(индекс, име и презиме, ниво студија)$
  - $Смер(ознака, назив, трајање)$
- које су у односу:
  - $Студент(*) \text{---} студира \text{---} (1) \text{---} Смер$ 
    - (ознаке кардиналности као у УМЛ-у)
- Све зависности у релацијама  $Студент$  и  $Смер$  су исправне...



...

- Редундантност може да се уочи, али је очигледнија када се у модел уведу и кључеви везаних релација (или се посматра међусобна зависност односа и атрибута):
  - $Студент(индекс, име и презиме, ниво студија, ознака смера)$
  - $Смер(ознака, назив, трајање)$
- Сада је очигледно да постоји редундантност атрибута  $ниво студија$  у релацији  $Студент$ 
  - Због зависности:  $ознака смера \rightarrow ниво студија$
- Исправно решење би било:
  - $Студент(индекс, име и презиме, ознака смера)$
  - $Смер(ознака, назив, ниво студија, трајање)$

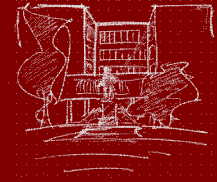


- Да би се редувантности виделе у концептуалном моделу, модел би морао да се непрекидно посматра критички и при томе део по део „преводи“ у релациони (тј. логички) модел
  - то није оствариво пре интегрисања погледа
    - превођење једног погледа може да да непотпун “логички” модел без довољно информација да би се уочиле редувантности
  - то се ретко ради и након интегрисања погледа
    - нема много смисла да се више пута прави логички модел и поправља концептуални, зато што редувантности често не представљају проблем у концептуалном моделу

## Пројектовање база података

Саша Малков

[P273]



Тема 8.5

## Логичко моделирање - Основне нормалне форме



## Нормалне форме

- *Нормалне форме* су специфични облици релација који задовољавају одређена правила у вези функционалних зависности
  - који зато *гарантују* да у релацији неће бити редувантности одређеног типа



## Нормалне форме

- У свом првом раду о релационом моделу (1970), Код је дефинисао само
  - 1. нормалну форму (1NF)
- Годину дана касније (1971) је дефинисао
  - 2. нормалну форму (2NF)
  - 3. нормалну форму (3NF)
- Заједно са Рејмондом Бојсом је дефинисао (1974)
  - Бојс-Кодову нормалну форму (BCNF)
- Касније су формулисане још и
  - 4. нормална форма (4NF) (1977)
  - 5. нормална форма (5NF) (1979)
  - Нормална форма домена и кључа (DKNF) (1981)
  - Нормална форма елементарног кључа (EKNF) (1982)
  - 6. нормална форма (6NF) (2003)
  - Нормална форма есенцијалних торки (ETNF) (2012)
  - Нормална форма наткључева (SKNF)
  - Нормална форма без редуванси (RFNF)



## Нормалне форме (2)

- Већина нормалних форми (све осим  $6NF$ ) могу да се поређају тако да наредна подразумева све претходне:
  - 1. нормална форма ( $1NF$ )
  - 2. нормална форма ( $2NF$ )
  - 3. нормална форма ( $3NF$ )
  - Нормална форма елементарног кључа ( $EKNF$ )
  - Бојс-Кодова нормална форма ( $BCNF$ )
  - 4. нормална форма ( $4NF$ )
  - Нормална форма есенцијалних торки ( $ETNF$ )
  - Нормална форма без редунданси ( $RFNF$ )
  - Нормална форма суперкључева ( $SKNF$ )
  - 5. нормална форма ( $5NF$ )
  - Нормална форма домена и кључа ( $DKNF$ )
- 6. нормална форма ( $6NF$ ) представља алтернативу за  $DKNF$  и не може да се непосредно упореди са осталим нормалним формама



## 1. нормална форма

- Релација је у првој нормалној форми ако сваки атрибут може да има само атомичне вредности.
  - Релациони модел претпоставља атомичне вредности
  - ЕР омогућава и сложене вредности атрибута



## 1. нормална форма (пример)

- Нека један предмет може да држи више наставника
- *Неатомични* модел би омогућавао да у релацији *џредмети* постоји атрибут *наставници* са именима више наставника
- $1NF$  то забрањује
- Једно могуће решење је да се атрибут *наставници* замени већим бројем атрибута *наставник1*, *наставник2*, *наставник3*
  - такво решење није добро ни ако има мање ни ако има више од 3 наставника за неки предмет, па чак ни када их је тачно 3
- Боље решење је да се уведе нова релација која моделира однос предмета и наставника



## 2. нормална форма

- Релација је у другој нормалној форми
  - ако је у првој нормалној форми и
  - ако ниједан некључан атрибут (тј. који не припада ниједном кандидату кључу) није функционално зависан од неког правог подскупа атрибута неког кандидат кључа.
- Суштина је да у случају сложеног кључа сви остали атрибути зависе од *целој* кључа а не од неког његовог дела.
  - т.ј. кандидати кључеви су минимални



## 2. нормална форма (пример)

- Нека следећа релација описује предмете које су уписали студенти:
  - УписанПредмет{ **индекс**, **шифр**ред, име, презиме, називред, којиуш }
  - Релација је у 1.НФ
- Релација није у 2.НФ
  - *име* и *презиме* зависе од атрибута *индекс*
  - *називред* зависи од атрибута *шифр*ред
- Да би се задовољила 2.НФ потребно је да се релација подели на више релација



## 2. нормална форма (пример)

- Релацију
  - УписанПредмет{ **индекс**, **шифр**ред, име, презиме, називред, којиуш }
- Можемо да поделимо на 3 релације:
  - УписанПредмет{ **индекс**, **шифр**ред, којиуш }
  - Студент{ **индекс**, име, презиме }
  - Предмет{ **шифр**ред, називред }
- Према теорему о декомпозицији, декомпозиција је потпуна
  - Показује се за једну по једну декомпозицију, најпре:
    - УписанПредмет{ **индекс**, **шифр**ред, називред, којиуш }
    - Студент{ **индекс**, име, презиме }
  - а затим и
    - УписанПредмет{ **индекс**, **шифр**ред, којиуш }
    - Предмет{ **шифр**ред, називред }



## 2. нормална форма (дискусија)

- Нарушавање 2.НФ указује на неисправно груписање атрибута у релације
  - релација моделира више различитих ФЗ
  - то води редувантности
- Постизање 2.НФ је кључно за позиционирање атрибута у *одговарајуће* релације
- Ипак, 2.НФ допушта да у релацији постоје два кључа кандидата, као и транзитивне зависности
- Увек је могућа потпуна декомпозиција до 2НФ



## 3. нормална форма

- Ако је  $R$  релација и  $X$  неки подскуп њених атрибута и  $A$  неки атрибут релације  $R$ , онда је  $R$  у 3.НФ ако за сваку ФЗ  $X \rightarrow A$  на релацији  $R$  важи једно од:
  - $A \in X$ , тј.  $A$  припада скупу  $X$  (тзв. тривијална зависност)
  - $X$  је наткључ (тј. кључ или садржи кључ)
  - $A$  је део неког кључа релације
- Оригинална дефиниција (Код, 1971)
  - Релација је у 3.НФ ако је
    - у 2.НФ и
    - сваки неключан атрибут је нетранзитивно зависан од сваког кључа.







## Бојс-Кодова нормална форма

- Ако је  $R$  релација и  $X$  неки подскуп њених атрибута и  $A$  неки атрибут релације  $R$ , онда је  $R$  у Бојс-Кодовој нормалној форми ако за сваку ФЗ  $X \rightarrow A$  на релацији  $R$  важи једно од:
  - $A \in X$ , тј.  $A$  припада скупу  $X$ , тзв. тривијална зависност
  - $X$  је наткључ
- Тј.
  - сваки атрибут релације је или део кључа или зависи од *цело*т кључа
  - не постоје зависности међу некључним атрибутима ни транзитивне зависности



## Бојс-Кодова нормална форма, дискусија

- Релације које су у 3.НФ најчешће задовољавају и БКНФ
  - Практично увек када постоји само један кључ или нема преклапања међу кључевима
  - Већина примера “релација у 3НФ које нису у БКНФ”, који се могу пронаћи у литератури, заправо нису исправни, тј. обично нису у 3НФ иако аутори тврде супротно
- Штавише, тешко је пронаћи добар пример који
  - јесте у 3.НФ
  - није у БКНФ
  - може да се преведе у БКНФ



## БКНФ, пример 1.1

- ([https://en.wikipedia.org/wiki/Boyce-Codd\\_normal\\_form](https://en.wikipedia.org/wiki/Boyce-Codd_normal_form))
- Посматрамо релацију
  - Резервација{ шерен, о~~г~~, го, шарифа }
    - постоје два терена, један је травнати и један бетонски
    - резервације су дефинисане тереном и периодом
    - постоје 4 тарифе – по две за сваки од терена зависно да ли је термин предвиђен за чланове или не
- Релација јесте у 2.НФ и 3.НФ (штавише, нема атрибута који су ван кључева)
  - постоји већи број кључева
    - { шерен, о~~г~~, го, шарифа }
    - { шерен, о~~г~~, го, шарифа }
    - { шерен, о~~г~~, го, шарифа }
    - { шерен, о~~г~~, го, шарифа }
- Релација *није* у БКНФ
  - проблем је у зависности шарифа -> шерен
    - шарифа није наткључ
    - шерен не припада скупу { шарифа }



## БКНФ, пример 1.2

- Једно решење је додавање новог атрибута за\_чланове:
  - Резервација{ шерен, о~~г~~, го, шарифа, за\_чланове }
- Сада постоје додатне Ф.З:
  - { шарифа } -> { шерен, за\_чланове }
  - { шерен, за\_чланове } -> { шарифа }
- Користимо другу за потпуну декомпозицију:
  - Резервација{ шерен, о~~г~~, го, за\_чланове }
  - Тарифа{ шерен, за\_чланове, шарифа }
    - може и обрнуто, да кључ буде шарифа, зато што је ФЗ у оба смера
- Испоставља се да је проблем у *скривеним* ФЗ које постају очигледне када се додају нови атрибути



## БКНФ, пример 2.1

- Посматрамо релацију
  - $Rasipored\{ \underline{радник}, \underline{тим}, \underline{руководилац} \}$
- и зависности
  - $радник, тим \rightarrow руководилац$ 
    - сваки радник може да буде у више тимова
    - у сваком тиму радник има тачно једног руководиоца
    - у једном тиму може да буде више руководилаца
      - али су одговарајући скупови подређених радника дисјунктни
  - $руководилац \rightarrow тим$ 
    - руководилац може да руководи само у једном тиму



## БКНФ, пример 2.2

- (Teorey, 2011) наводи овај пример:
  - “релација јесте у 3.НФ а није у БКНФ”
  - проблем је у аномалији брисања – ако сви радници код неког руководиоца напусте тим, онда ће се изгубити информација о руководиоцу у тиму
  - предлаже да се проблем реши тако да се додатно уведе релација  $Руководилац\{ \underline{руководилац}, \underline{тим} \}$  која би описивала другу зависност



## БКНФ, пример 2.3

- Претходни проблем се делимично решава ако уочимо да имамо још једну зависност:
  - $радник, руководилац \rightarrow тим$
  - ова зависност не изгледа *природно*, тим пре што може да се изведе из постојећих, али је она заправо непосредна последица
    - ограничења да сваки руководилац руководи највише једним тимом
    - и претпоставке да сваки радник има руководиоца у сваком тиму у који је распоређен
- Ако ту ФЗ прогласимо за основну, а раније установљену  $радник, тим \rightarrow руководилац$  за изведену...
- ...онда видимо да није задовољена чак ни 2.НФ (а самим тим ни 3.НФ):
  - $радник, руководилац \rightarrow руководилац \rightarrow тим$



## БКНФ, пример 2.4

- Да бисмо задовољили 2.НФ морамо да направимо две релације:
  - $Rasipored\{ \underline{радник}, \underline{руководилац} \}$
  - $Руководилац\{ \underline{руководилац}, \underline{тим} \}$
- Сада су задовољене и 2.НФ и 3.НФ и БКНФ
- Решење је једноставније од изворног
  - у релацији  $Rasipored$  нема атрибута  $тим$
  - ...али је изгубљено ограничење да у сваком тиму радник има тачно једног руководиоца
  - ...зато што смо напустили првобитну ФЗ
- Видимо да полазни пример није задовољавао 2.НФ (а тиме ни 3.НФ) зато што нису довољно пажљиво разматране функционалне зависности
- Али видимо и да када додајемо нове зависности не смемо да занемаримо старе, иначе можемо да их изгубимо



## БКНФ и НФЕК

- Нека имамо релацију:
  - ПријаваИспитиџа ( стџудентџ, исџиџни рок, џредметџ )
  - кључ је { стџудентџ, исџиџни рок }
- и зависности:
  - { стџудентџ, исџиџни рок } → { џредметџ }
    - тј. студент у једном исп. року може да пријави испит из највише једног предмета
  - { џредметџ } → { исџиџни рок }
    - тј. сваки предмет се полаже само у по једном испитном року
- Тада:
  - релација ПријаваИспитиџа није у БКНФ:
    - испитни рок зависи од предмета, а предмет није део кључа
    - ако направимо декомпозицију, нарушићемо прву зависност...
    - ако променимо кључ ( стџудентџ, џредметџ ), проблем опстаје



## БКНФ и НФЕК (2)

- Али:
  - релација ПријаваИспитиџа **јесџе** у НФЕК
  - прва зависност је од целог кључа
  - друга је према делу кључа
  - обе су елементарне



## БКНФ и НФЕК (3)

- Проблем не може добро да се реши као у претходном примеру, увођењем алтернативних зависности
- Постоји потпуна декомпозиција:
  - ПријаваИспитиџа ( стџудентџ, џредметџ )
  - ПредметџУРоку ( џредметџ, исџиџни рок )
- Она је потпуна, али **не чува све зависности**
  - није очувано ограничење да у једном испитном року студент може да пријави испит из највише једног предмета
- Штавише, не постоји декомпозиција која чува све ФЗ



## БКНФ и НФЕК (4)

- Значај НФЕК је у томе што неке релације не могу да пређу пут од ЗНФ до БКНФ а да се очувају све зависности...
- ...али могу да дођу бар до НФЕК
- Да би се у моделу сачувале све зависности, потребно је да се постојећој релацији дода нова релација која описује другу зависност:
  - ПредметџУРоку ( џредметџ, исџиџни рок )
  - Уводи се редундантност...
    - ...која се решава додавањем страног кључа (предмет, испитни рок) из релације ПријаваИспитиџа према релацији ПредметџУРоку
    - (што не подржавају сви РСУБП, зато што се страни кључ не везује за примарни)

## Литература за тему



- Teorey, Lightstone, Nadeau, Jagadish, **Database Modeling and Design**, 5.ed, Elsevier, 2011.
- Watt, Eng, **Database Design**, 2.ed, Open Edition, 2014.
- Гордана Павловић-Лажетић, **Увод у релационе базе података**, 2.изг. *Математички факултет*, 1999.
  - доступно онлајн: <http://poincare.matf.bg.ac.rs/~gordana/urbp-2016.htm>
- Ramakrishnan, Gehrke, **Database Management Systems**, 2.ed, 2000.
- Darween, Date, Fagin, **A Normal Form for Preventing Redundant Tuples in Relational Databases**, *ICDT '12 Proceedings*, 2012.