

Normální formy

Zdeněk Kouba

Příklad špatného modelu

Osob Cislo	Prijmeni	Cislo Kancelare	Obec	Ulice	CP	PSC	Kraj	Hejtman
1001	Novák	238	Liteň	Hlavní	10	26727	Středočeský	Rath
1001	Novák	238	Bystřice	Benešova	112	25751	Středočeský	Rath
1001	Novák	238	Benešov	Jiráskova	124	25601	Středočeský	Rath
1201	Řezáč	204	Praha	Ječná	21	12000	Praha	Bém
1221	Falc	216	Kladno	Uhelná	234	27201	Středočeský	Rath
1234	Lupínek	308	Kladno	Rudná	23	27202	Středočeský	Rath

Paradox při update:

- Když pan Novák změní číslo kanceláře, musíme tuto změnu provést ve všech jeho záznamech (pro všechny jeho adresy).
- Podobně změna hejtmana Středočeského kraje by vyvolala změnu ve všech záznamech osob s bydlištěm ve Středočeském kraji.

Příklad špatného modelu

Osob Cislo	Prijmeni	Cislo Kancelare	Obec	Ulice	CP	PSC	Kraj	Hejtman
1001	Novák	238	Liteň	Hlavní	10	26727	Středočeský	Rath
1001	Novák	238	Bystřice	Benešova	112	25751	Středočeský	Rath
1001	Novák	238	Benešov	Jiráskova	124	25601	Středočeský	Rath
1201	Řezáč	204	Praha	Ječná	21	12000	Praha	Bém
1221	Falc	216	Kladno	Uhelná	234	27201	Středočeský	Rath
1234	Lupínek	308	Kladno	Rudná	23	27202	Středočeský	Rath

Paradox při insertu:

- Když budeme vkládat záznam nového zaměstnance pana Jonáše s bydlištěm v Jihočeském kraji, musíme si nejprve zjistit jméno hejtmána Jihočeského kraje, abychom mohli tuto informaci zadat spolu s adresou pana Jonáše.

Příklad špatného modelu

Osob Cislo	Prijmeni	Cislo Kancelare	Obec	Ulice	CP	PSC	Kraj	Hejtman
1001	Novák	238	Liteň	Hlavní	10	26727	Středočeský	Rath
1001	Novák	238	Bystřice	Benešova	112	25751	Středočeský	Rath
1001	Novák	238	Benešov	Jiráskova	124	25601	Středočeský	Rath
1201	Řezáč	204	Praha	Ječná	21	12000	Praha	Bém
1221	Falc	216	Kladno	Uhelná	234	27201	Středočeský	Rath
1234	Lupínek	308	Kladno	Rudná	23	27202	Středočeský	Rath

Paradox při delete:

- Zrušíme-li po odchodu pana Řezáče z firmy jeho záznam, ztratíme současně s tím informaci o tom, že primátorem (hejtmanem) Prahy je pan Bém.

Konvence

Nechť $X = \{X_1, X_2, \dots, X_n\}$. Symbolem X^\times označme kartézský součin

$$X^\times = X_1 \times X_2 \times \dots \times X_n .$$

Symbol x^\times pak bude označovat nějakou uspořádanou entici $x^\times \in X^\times$.

Funkční závislost

Mějme schema $R(A)$, kde A je množina všech atributů relace R .

Mějme množiny atributů $X, Y \subseteq A$.

Říkáme, že Y funkčně závisí na X právě tehdy, když pro každou hodnotu $x^x \in X^x$ existuje nejvýše jedna hodnota $y^x \in Y^x$.

Funkční závislost značíme $X \rightarrow Y$.

Nezávisí-li Y na X , značíme to symbolem $X \nrightarrow Y$.

Funkční závislost - příklad

Mějme relaci:

AdresaOsoby = {OsobCislo, Prijmeni, CisloKancelare,
Obec, Ulice, CP, PSC, Kraj, Hejtman}

Příklady funkčních závislostí:

{OsobCislo} → {OsobCislo, Prijmeni, CisloKancelare,
Obec, Ulice, CP, PSC, Kraj, Hejtman}

{Kraj, Obec, Ulice} → {PSC}

{Kraj} → {Hejtman}

Parciální funkční závislost

Mějme množiny atributů X a Y .

Nechť $X \rightarrow Y$ (t.j. Y je funkčně závislé na X).

$X \rightarrow Y$ nazveme **parciální** funkční závislostí právě tehdy, existuje-li nějaká množina atributů $X_1 \subset X$ taková, že $X_1 \rightarrow Y$.

Parciální funkční závislost - příklad

Mějme schema $R(A)$, kde A je množina všech atributů relace R .

$A = \{\text{Predmet, Ucitel, Termin, Mistnost, Kredity}\}$

Funkční závislost $\{\text{Predmet, Termin}\} \rightarrow \{\text{Kredity}\}$
je parciální funkční závislostí, neboť

$\{\text{Predmet}\} \rightarrow \{\text{Kredity}\}$

a přitom

$\{\text{Predmet}\} \subset \{\text{Predmet, Termin}\}$

Transitivní funkční závislost

Mějme množiny atributů X a Y , dále mějme atribut C takový, že $C \notin X$ a $C \notin Y$.

Říkáme, že atribut C je transitivně závislý na X právě tehdy, když současně platí:

- $X \rightarrow Y \rightarrow C$
- $Y \not\rightarrow X$.

Transitivní funkční závislost - příklad

Mějme schema $R(A)$, kde A je množina všech atributů relace R .

$A = \{\text{OsobCislo, Obec, Ulice, CP, PSC, Kraj, Hejtman}\}$

$\{\text{OsobCislo, Obec, Ulice, CP, PSC}\} \rightarrow \{\text{Kraj}\}$

$\{\text{Kraj}\} \rightarrow \{\text{Hejtman}\}$

$\{\text{Kraj}\} \not\rightarrow \{\text{OsobCislo, Obec, Ulice, CP, PSC}\}$

Tedy:

Množina atributů $\{\text{Hejtman}\}$ je transitivně závislá na množině atributů $\{\text{OsobCislo, Obec, Ulice, CP, PSC}\}$

První normální forma

Relace je v první normální formě právě tehdy, když platí současně:

1. Atributy atomické (dále nedělitelné)
2. Přístup k řádkům relace podle obsahu (klíčových) atributů
3. Jedinečné m-tice (řádky)

Příklad relace, která není v 1.NF:

Rodiče		
Otec	Matka	Děti
Jan	Marta	{Eliška, Lucie}
Jan	Marie	{Jiří, Josef}
Josef	Marta	{Leoš}

Druhá normální forma

Relace je ve druhé normální formě právě tehdy, když platí současně:

1. Je v 1. NF.
2. Žádný neklíčový atribut není parciálně funkčně závislý na žádném z klíčů dané relace.

Například relace {IdStudenta, IdPredmetu, JmenoStudenta, Semestr} není ve 2NF, neboť:

- Klíčem je {IdStudenta, IdPredmetu}
- Neklíčový atribut JmenoStudenta je funkčně závislý na {IdStudenta}, je tedy parciálně závislý na klíči.
- Neklíčový atribut Semestr je funkčně závislý na {IdPredmetu}, je tedy parciálně závislý na klíči.

Druhá normální forma

Například relace $\{\underline{\text{IdStudenta}}, \underline{\text{IdPredmetu}}, \text{JmenoStudenta}, \text{Semestr}\}$ není ve 2NF, neboť:

- Klíčem je $\{\underline{\text{IdStudenta}}, \underline{\text{IdPredmetu}}\}$
- Neklíčový atribut JmenoStudenta je funkčně závislý na $\{\underline{\text{IdStudenta}}\}$, je tedy parciálně závislý na klíči.
- Neklíčový atribut Semestr je funkčně závislý na $\{\underline{\text{IdPredmetu}}\}$, je tedy parciálně závislý na klíči.

Řešení:

Relaci rozdělíme do tří tabulek

- $\{\underline{\text{IdStudenta}}, \underline{\text{IdPredmetu}}\}$
- $\{\underline{\text{IdStudenta}}, \text{JmenoStudenta}\}$
- $\{\underline{\text{IdPredmetu}}, \text{Semestr}\}$

Třetí normální forma

Relace je ve třetí normální formě právě tehdy, když platí současně:

1. Je v 2. NF.
2. Žádný neklíčový atribut není transitivně funkčně závislý na žádném z klíčů dané relace.

Například relace {IdStudenta, JmenoStudenta, Fakulta, Dekan} není ve 3NF, neboť:

- Sice je ve 2NF, ale
- Atribut Dekan je funkčně závislý na {Fakulta}
- {Fakulta} je funkčně závislá na {IdStudenta} (předpokládáme, že student nemůže být současně studentem více fakult téže university).
- {IdStudenta} není funkčně závislá na {Fakulta}.

Atribut Dekat je tedy transitivně funkčně závislý na klíči.

Třetí normální forma

Například relace {IdStudenta, JmenoStudenta, Fakulta, Dekan} není ve 3NF, neboť atribut Dekat je transitivně funkčně závislý na klíči.

Řešení:

Relaci rozdělíme do tabulek

- {IdStudenta, JmenoStudenta, Fakulta}
- {Fakulta, Dekan}