



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



Relační datový model I.

Ing. Zbyněk Pospěch

Střední průmyslová škola, Obchodní akademie
a Jazyková škola s právem státní jazykové zkoušky, ve Frýdku-Místku

Kód DUM
II-S1-05

Základní informace

II-S1-05

Název: Relační datový model I.

Autor: Ing. Zbyněk Pospěch

Anotace: výukový materiál objasňuje původ do relačního systému řízení báze dat, relační datový model, terminologie relačního modelu, vlastnosti relace, integritní omezení, schéma a instance relační databáze, hodnota NULL, super klíč, kandidátní klíč a primární klíč.

Studijní obor: Informační technologie

Ročník: I.

Předmět: Relační databázové systémy

Forma DUM: Prezentace

Datum vytvoření: 17.10. 2012

Datum aktualizace: 21.6.2013

Předseda komise: Ing. Marta Murínová

Metodik DUM: Ing. Roman Wyka

Jazyková korekce: Mgr. Jarmila Kotásková

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Relační systém řízení báze dat (RSŘBD)

- ▶ Mezi používanými SŘBD v současnosti převládá.
- ▶ Je založen na relačním modelu dat navrženém E. F. Codd v roce 1970.
- ▶ Cíle relačního modelu:
 - vysoký stupeň nezávislosti dat
 - zabezpečení sémantiky, konzistence a redundance dat pomocí tzv. „normalizovaných relací“
 - využití množinově orientovaných jazyků pro manipulaci s daty

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Relační datový model

– formální zápis

- ▶ relace R nad množinou atributů $A = \{A_1, A_2 \dots A_n\}$
- ▶ jména atributů $A_1, A_2 \dots, A_n$
- ▶ domény atributů $D_i = \text{dom}(A_i)$
- ▶ jméno relace R
- ▶ schéma relace $R(A_1:D_1, A_2:D_2, \dots A_n:D_n)$ nebo zkráceně $R(A)$
- ▶ hodnoty atributů jsou atomické \Rightarrow 1. normální forma (1NF) viz dále
- ▶ relace $R^*(A) \subset D_1 \times D_2 \times \dots \times D_n$
- ▶ n -tice $(a_1, a_2 \dots a_n)$

Relační datový model – příklad

Žáci					
Rodné číslo	Jméno	Příjmení	Datum přijetí	Třída	Ročník
9002011234	Jan	Nový	1.9.2012	A	1
9105035555	Petr	Tichý	10.9.2012	B	1
9061230001	Jana	Malá	1.9.2011	A	2

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Relační datový model – příklad

- ▶ **Jméno relace:**
 - Žáci
- ▶ **Jména atributů:**
 - Jméno, Příjmení, Datum narození, Třída, Ročník.
- ▶ **Domény atributů:**
 - Řetězec znaků → Jméno, Příjmení a Třída, Rodné číslo (není číslo),
 - Celé čísla → Ročník,
 - Datum → Datum přijetí.
- ▶ **Schéma relace:**
 - Žáci (Jméno: Řetězec znaků, Příjmení: Řetězec znaků, Datum přijetí: Datum, Třída: Řetězec znaků , Ročník: Celá čísla).
- ▶ **Datové n-tice:**
 - (9002011234, Jan, Nový, 1.9.2012, A, 1);
 - (9105035555, Petr, Tichý, 10.9.2012, B, 1);
 - (9061230001, Jana, Malá, 1.9.2011, A, 2).

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Vlastnosti relace

- ▶ Relace má jedinečné jméno, jež ji odlišuje od jiných relací v databázi.
- ▶ Každý atribut musí mít v relaci jedinečné jméno.
- ▶ Datové n-tice mohou být v libovolném pořadí.
Poznámka: Prakticky pořadí může ovlivnit efektivitu přístupu.
- ▶ Atributy mohou být v libovolném pořadí.
- ▶ Všechny atributy v relaci musí být elementární – nedělitelné – 1.NF .
- ▶ Každá datová n-tice musí být jedinečná, neexistují duplicitní datové n-tice.
- ▶ Všechny hodnoty v daného atributu musí být ze stejné domény.

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Integritní omezení (Relační integrita)

- ▶ V relaci existuje omezená množina přípustných hodnot daných atributů.
- ▶ Množina přípustných hodnot je definována doménou a integritním omezením (I) .
- ▶ **Příklady integritního omezení:**
 - IO1 – Každý žák má jedinečné rodné číslo.
 - IO2 – Žáci mohou navštěvovat pouze třídu A nebo B, neboli atribut **Třída**=**{ A, B }**
 - IO3 – Žáci mohou navštěvovat pouze **Ročník**=**{ 1, 2, 3, 4, 5 }**

Úplná definice integritních omezení je vysvětlena ve výukovém materiálu číslo II-S1-06.

Schéma a instance relační databáze

- ▶ Úplná definice schématu relace = **schéma relační databáze (R,I)**
 - $R = \{ R_1, R_2, \dots R_k \}$ – všechny relace v databázi
 - I – množina integritních omezení
- ▶ Množina všech přípustných relačních databází se schématem (R,I) je množina relací $R_1^*, R_2^*, \dots R_k^*$ takových, že jejich prvky vyhovují I .
- ▶ Instance relačního databázového schématu
 - Je jedinečný výskyt relační databáze odpovídající danému relačnímu databázovému schématu (R,I) .

Terminologie relačního modelu

- ▶ Terminologie relačního modelu je poměrně složitá, dána vícevrstvou architekturou, pro každou fázi návrhu jsou definovány obsahově jiné termíny.
- ▶ Relační datový model má následující hlavní složky.

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Terminologie relační datové struktury dle modelu.

Realita	Konceptuální model (E-R model)	Logická úroveň – Relační datový model	Fyzická úroveň – SQL
Objekt	Entita	Relace	Tabulka
Vlastnost objektu	Atribut	Atribut	Pole (sloupec)
Množina přípustných hodnot pro jednu nebo více vlastností	V širším kontextu to jsou všechny hodnoty odpovídající pravidlům integritních relačních schémat.		
Typ objektu	Schéma entity	Schéma relace	Struktura tabulky
Instance objektu	Výskyt entity	Datová n-tice	Záznam (řádek)
Skupina objektů, jež jsou ve vzájemných vztazích	Model	Relační databáze	Kolekce normalizovaných tabulek
Vztahy mezi objekty	Relace	Vztah	Relace

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Hodnoty NULL

- ▶ Jedná se o hodnotu atributu, která není definována.
- ▶ V praxi se jedná o hodnotu nezadanou nebo neznámou.
- ▶ Hodnota NULL musí být v definici doménového omezení daného atributu povolena. Pokud v daném atributu není hodnota NULL povolena, je povinnost zadat u daného atributu hodnotu ve všech datových n-ticích.

NULL není nula



- ▶ Příklad chybné záměny hodnoty NULL za 0: Žák číslo 1 nepsal test číslo 5 a 3, byla mu chybně zadána známka 0.
- ▶ Výsledný průměr známek je $(5+0+3+0)/4=2$!!!

Chybně zadáné hodnoty NULL		
ID_testu	ID_žáka	Známka
1	1	5
5	1	0
4	1	3
3	1	0

Správně zadáné hodnoty NULL		
ID_testu	ID_žáka	Známka
1	1	5
5	1	.NULL.
4	1	3
3	1	.NULL.

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Super klíč a kandidátní klíč

► Super klíč

- atribut nebo množina atributů jednoznačně identifikující datovou n-tici v relaci
- Může obsahovat nadbytečné atributy, které pro jednoznačné určení datové n-tice nejsou potřebné.

► Kandidátní klíč

- Super klíč, který obsahuje minimální počet atributů k jedinečné identifikaci záznamů.
- **vlastnosti kandidátního klíče:**
 - **jedinečnost** – jednoznačně identifikující datovou n-tici v relaci
 - **neredukovatelnost** – žádná vlastní podmnožina atributů kandidátního klíče nezajišťuje jedinečné určení záznamů

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Příklad superklíčů a kandidátního klíče

- Předpoklad relace Znamky: V daném ročníku a pololetí přidělí jeden učitel jednomu žáku v jeden den z jednoho předmětu z jednoho tématu jednu známku.

Relace Znamky

- RČ_žák
- Předmět
- RČ_Učitel
- Ročník
- Pololetí
- Téma
- Váha_znamky
- Datum
- Klasifikace
- Poznámky

Superklíč 1.

- RČ_žák
- Předmět
- RČ_Učitel
- Ročník
- Pololetí
- Téma
- Váha_znamky
- Datum
- Klasifikace
- Poznámky

Superklíč 2.

- RČ_žák
- Předmět
- RČ_Učitel
- Ročník
- Pololetí
- Téma
- Váha_znamky
- Datum
- Klasifikace

Superklíč 3.

- RČ_žák
- Předmět
- RČ_Učitel
- Ročník
- Pololetí
- Téma
- Váha_znamky
- Datum

Superklíč 4.

- RČ_žák
- Předmět
- RČ_Učitel
- Ročník
- Pololetí
- Téma
- Datum
- Klasifikace

Kandidátní klíč

- RČ_žák
- Předmět
- RČ_Učitel
- Ročník
- Pololetí
- Téma
- Datum

Primární klíč

- ▶ Kandidátní klíč, který je vybrán, aby jedinečně určoval záznamy v tabulce.
- ▶ Klíč K schématu $R(A)$ je minimální množina atributů z A , jejichž hodnoty jednoznačně určují prvky z R^* .
- ▶ Je-li K klíč schématu $R(A)$, pak pro každou přípustnou relaci R^* platí:
Jsou-li a a b dvě různé n -tice z R^* , pak existuje alespoň jeden atribut $A_i \in K$ takový, že $a[K] \neq b[K]$.
- ▶ Každá relace má definován právě jediný primární klíč.

Primární klíč (PK)

- ▶ Druhy primárního klíče dle počtu atributů
 1. Jednoduchý primární klíč – PK je založen na jediném atributu.
 2. Složený primární klíč – PK je založen na dvou nebo více attributech.

Přirozený primární klíč

- ▶ PK je nastaven nad atributem nebo atributy, které přirozeně jednoznačně identifikují datovou n-tici relace.

Entita	Atribut přirozeného primárního klíče
Osoba	Rodné číslo
Faktura	Číslo faktury
Majetek	Inventární číslo majetku
Známka	Složený klíč obsahuje příliš mnoho atributů: RČ_Žák + Předmět + RČ_Učitel + Téma + Váha_známky + Ročník + Pololetí + Datum, vhodnější bude klíč syntetický ID_ZNAMKY

Syntetický (umělý) primární klíč

- ▶ Používá se, pokud neexistuje přirozený atribut PK nebo se jedná o složený PK, který je příliš složitý a jehož použití je nepraktické.
- ▶ PK založen na umělém atributu obsahující pořadové nebo pseudonáhodné jedinečné číslo.
- ▶ Obvykle se označuje jako **IDENTIFIKÁTOR**, ve jménu atributu začínající zkratkou **ID_**

Otázky k opakování



- ▶ Teoretické otázky:
 1. Vysvětlete strukturu relačního datového modelu.
 2. Vyjmenujte vlastnosti relace.
 3. Definujte schéma relační databáze.
 4. Vysvětlete hodnotu NULL a uveďte příklady.
 5. Co je to integritní omezení ? Uveďte příklad.
 6. Co jsou to klíče, jaké klíče znáte a jak jsou definovány ? Uveďte příklady klíčů.

Otázky k opakování



► Praktický úkol:

1. Navrhněte tři nezávislé relace **Faktury**, **Žáci** a **Katalog_Zboží**,
2. Určete atributy relace a domény atributů.
3. Zajistěte, aby relace odpovídaly 1. normální formě.
4. Zajistěte entitní a doménovou integritu.
5. Zvolte vhodně pro atributy relace povolení nebo zakázání hodnot NULL.
6. Určete superklíče, kandidátní a primární klíče relace.

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Informační zdroje

1. CONOLLY, Thomas, Carolyn E BEGG a Richard HOLOWCZAK. *Mistrovství – databáze: profesionální průvodce tvorbou efektivních databází*. Vyd. 1. Brno: Computer Press, 2009, 584 s. ISBN 978–80–251–2328–7.
2. DUBOIS, Paul. *MySQL profesionálně: komplexní průvodce použitím, programováním a správou MySQL*. Vyd. 1. Překlad Jan Pokorný. Brno: Mobil Media, 2003, 1071 s. ISBN 80–865–9341–X.
3. TELNAROVÁ, Zdeňka. OSTRAVSKÁ UNIVERZITA. *Relační databáze: Distanční výuková podpora*. Aktualizovaná verze 2006. Ostrava: pouze v elektronické formě, 2006. ISBN (bez uvození).

Citace obrázků

- ▶ Všechny obrázky jsou vytvořeny autorem prezentace, 2012.
- ▶ Ve schématech využil autor ikony z galerie Microsoft Office 2010.

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

Děkuji za pozornost

»» Ing. Zbyněk Pospěch