

## 33. Relační datový model a jazyk SQL.

**Relační model dat** – spadá mezi tzv. databázové modely – zahrnuje:

- Definici logické struktury dat v relační databázi, tj. říká, jak jsou data v relační databázi strukturována na logické úrovni bez ohledu na vlastní implementaci.
- Definici obecných integritních pravidel, tj. omezení kladená na data, která musí platit v každé relační databázi
- Formální dotazovací jazyk (např. SQL), který umožňuje formulovat dotazy nad daty uloženými v relační databázi

Struktura relační databáze

- data jsou strukturována do tabulek (správně Relací)
- atributy - vlastnosti reálného objektu. Každý atribut může nabívat hodnoty z množiny přípustných hodnot - domény (často realizovanou pouze použitým datovým typem)
- domény mohou obsahovat pouze atomické hodnoty (žádné složené atributy)
- řádek tabulky (1 entita) dává do relace prvky jednotlivých domén a nazývá se n-tice
- aktivní doména - množina hodnot domény, které se aktuálně používají (doména=všechny možné hodnoty atributu, aktivní doména=aktuálně použité hodnoty atributů)

Databázová relace

Nechť  $D_1, D_2, \dots, D_n$  jsou množiny atomických hodnot označované jako domény.

Databázová relace na doménách  $D_1, D_2, \dots, D_n$  je dvojice  $R = (R, R^*)$ ,

- kde  $R = R(A_1:D_1, A_2:D_2, \dots, A_n:D_n)$  je schéma relace,
- kde  $A_i$  ( $A_i \neq A_j$  pro  $i \neq j$ ) značí jméno atributu definovaného na doméně  $D_i$ ,
- a  $R^* \subseteq D_1 \times D_2 \times \dots \times D_n$  je tělo relace.

Počet atributů  $n$  relace se označuje *stupeň (řád) relace*, kardinalita těla relace  $m = |R^*|$  se označuje *kardinalita relace*.

*citace: "Relační databáze se nazývají relační, protože data v nich jsou organizována v souladu s relačním modelem dat a ten se nazývá relační, protože data jsou podle tohoto modelu logicky strukturována na základě relací (tabulek). V některé literatuře se můžete dočíst, že se tyto databáze nazývají relační proto, že v nich existují relace mezi tabulkami. Toto není ten důvod. My už víme, že relace v relačním modelu označuje vlastně tabulku."*

### Integritní omezení

-jsou omezení kladená na data uložená v databázi

- obecná-musí platit v každé relační databázi bez ohledu na konkrétní aplikační zaměření.
- Specifická - jsou omezení specifická pro každou konkrétní aplikační oblast

### Kandidátní klíč (candidate key - CK)

atribut relace  $R$  se nazývá CK pokud splňuje (v každý okamžik):

- hodnoty atributu CK v relaci  $R$  jsou unikátní (jednoznačné)
- atribut CK je minimální v  $R$ . (je-li to atribut složený, tak z něho nelze vypustit žádnou

složku, aby nedošlo ke ztrátě unikátnosti)

### Primární klíč

- je jeden vybraný kandidátní klíč. Zbývajícím CK se říká *alternativní* nebo *sekundární*.
- u žádné složky primárního klíče nesmí chybět hodnota (NULL)

### Cizí klíč (foreign key FK)

je atribut relace R, když splňuje (vždy):

- Každá z hodnot FK je buď plně zadaná nebo plně nezadaná
- Existuje relace R1 s CK takovým, že každá zadaná hodnota FK je identická s hodnotou CK nějaké n-tice v R1

Hodnota FK může být (na rozdíl od PK!) nezadaná.

*Referenční integrita* - soulad hodnot cizího(FK) a odkazovaného klíče(CK)

### Pravidlo referenční integrity

Relační databáze nesmí obsahovat žádnou nesouhlasnou hodnotu cizího klíče

### Schéma relační databáze

nazýváme tak dvojici (R, I) kde:

- $R = \{R1, R2, \dots, Rk\}$  je množina schémat relací
- $I = \{I1, I2, \dots, Il\}$  je množina integritních omezení

## Relační algebra

Relační algebra je uzavřená. Tzn. jejími operandy jsou tabulky, stejně tak tabulku vrací. Operace jsou buď unární (nad jednou tabulkou) nebo binární (nad dvěma tabulkami). Výsledkem je vždy pouze jedna tabulka.

Relační algebrou rozumíme dvojici  $RA = (R, O)$ , kde nosičem R je množina relací a O je množina operací, která zahrnuje:

- tradiční množinové operace (**sjednocení, průnik, rozdíl, součin**)
  - sjednocení, průnik a rozdíl lze provádět pouze nad tabulkami se stejným schématem (stejnými sloupci)
  - kartézský součin** lze provádět i nad tabulkami s různým schématem (schéma obou relací se sloučí a nad těly relací se provede kartézský součin)
- Speciální relační operace
  - Projekce**-unární operace-výsledkem je tabulka, obsahující jenom některé sloupce původní
  - selektce (resektce)**- unární operace-výsledkem je tabulka obsahující všechny sloupce, ale jen některé řádky (ty vyhovující zadané podmínce)
  - spojení (přirozené)**- binární operace-výsledkem je tabulka obsahující schéma s atributy obou tabulek. V těle jsou n-tice, které vznikly spojením těch řádků tabulek, které měly stejné hodnoty ve stejné pojmenovaných sloupcích

*Minimální množina operací relační algebry* jsou sjednocení, rozdíl, kartézský součin, projekce a selekce.

Význam relační algebry

- je referenčním prostředkem pro hodnocení vlastností a porovnávání relačních dotazovacích jazyků
- je vhodným základem pro optimalizaci zpracování dotazů

## Relační kalkuly

jedná se o dotazovací jazyky na bázi logiky. Obsahuje formule logiky 1. řádu, tj. takové, ve kterých se vyskytují proměnné, konstanty, operátory porovnání, operátor negace, logické spojky a kvantifikátory.

Podle toho, zda jsou proměnné definovány na řádcích tabulek (n-ticích relací), nebo hodnotách ve sloupcích (hodnotách domény atributu) se rozlišují dva typy relačního kalkulu:

- n-ticový relační kalkul
- doménový relační kalkul

## Jazyk SQL

**SQL** (někdy vyslovováno anglicky *es-kjü-el* [ɛs kjʊː ɛl] [IPA](#), někdy též *s/kv* [siːkwəl] [IPA](#)) je standardizovaný **dotazovací jazyk** používaný pro práci s daty v relačních databázích. SQL je zkratka anglických slov **Structured Query Language** (strukturovaný dotazovací jazyk).

SQL příkazy se dělí na čtyři základní skupiny:

- Příkazy pro manipulaci s daty (**SELECT**, **INSERT**, **UPDATE**, **DELETE**, ...)
- Příkazy pro definici dat (**CREATE**, **ALTER**, **DROP**, ...)
- Příkazy pro řízení přístupových práv (**GRANT**, **REVOKE**)
- Příkazy pro řízení transakcí (**START TRANSACTION**, **COMMIT**, **ROLLBACK**)