

Logické datové modely

<small>Z ČWUT</small>

Obsah

- 1 Relační model (RMD)
 - 1.1 Relace
 - 1.2 Relační algebra
 - 1.3 Normální formy
 - 1.4 Algoritmy návrhu relačního schématu
- 2 Návrh relačního schématu ve 3. NF dekompozicí z univerzální relace
- 3 Návrh relačního schématu syntézou dle Bernsteina
- 4 Objektový model
- 5 Funkční závislosti

Relační model (RMD)

- RMD důsledně odděluje data, která jsou chápána jako relace, od jejich implementace
- přístup k datům je symetrický, tj. při manipulacích s daty se nezajímáme o přístupové mechanismy k datům obsažených v relacích
- pro manipulaci s daty jsou k dispozici dva silné prostředky - relační kalkul a algebra
- pro omezení redundance dat v relační databázi jsou k dispozici pojmy umožňující normalizovat relace, tj. vhodně navrhovat potřebné relační datové struktury

Relace

Relace vychází z matematického pojmu relace. Jde o množinu prvků, které mají tvar (a_1, a_2, \dots, a_n) , kde n je řád relace. Nazýváme je n -tice. Jednotlivé komponenty a_i jsou z domény atributů A_i . Doména je specifikovaná množina hodnot, kterých může atribut nabývat. $A_i : \text{dom}(A_i)$. Pojmenovaná relace $R, R(a_1:D_1, a_2:D_2, \dots, a_n:D_n) =$ Schéma relace.

Relační algebra

Relační algebra je množina operací, jejichž aplikace na nějaké relace vrací opět relaci. Protože relace jsou vlastně množiny lze využít bez zbytku množinové operace: součin, sjednocení, průnik, rozdíl. Součin lze použít vždy, pro ostatní operace musí mít relace stejný počet atributů.

Normální formy

- 0NF (nultá normální forma): Tabulka je v nulté normální formě právě tehdy, existuje-li alespoň jedno pole, které obsahuje více než jednu hodnotu.
- 1NF (první normální forma): Tabulka je v první normální formě, jestliže lze do každého pole dosadit pouze jednoduchý datový typ (jsou dále nedělitelné).
- 2NF (druhá normální forma): Tabulka je ve druhé normální formě, jestliže je v první a navíc platí, že existuje klíč a všechna neklíčová pole jsou funkcí celého klíče (a tedy ne jen jeho částí).
- 3NF (třetí normální forma): Tabulka je ve třetí normální formě, jestliže každý neklíčový atribut není transitivně závislý na žádném klíči schématu neboli je-li ve druhé normální formě a zároveň neexistuje jediná závislost neklíčových sloupců tabulky.

- BCNF (Boyce-Coddova normální forma): Tabulka je v Boyce-Coddově normální formě, jestliže pro každou netriviální závislost $X \twoheadrightarrow Y$ platí, že X obsahuje klíč schématu R .
- 4NF (čtvrtá normální forma): Tabulka je ve čtvrté normální formě, je-li ve třetí a popisuje pouze příčinnou souvislost (jeden fakt).
- 5NF (pátá normální forma): Tabulka je v páté normální formě, pokud je ve čtvrté a není možné do ní přidat nový sloupec (skupinu sloupců) tak, aby se vlivem skrytých závislostí rozpadla na několik dílčích tabulek.

Více na <http://home.zcu.cz/~kotouc/vyuka/dbm1/pr08.htm>

Algoritmy návrhu relačního schématu

Dva přístupy:

- normalizační teorie - návrh pomocí funkčních závislostí - dekompozice
- z konceptuálního schématu - pomocí transformačních pravidel

Aby byla možná dekompozice, musíme mít splněny 2 předpoklady:

- předpoklad schématu univerzální relace – zdůrazňuje jednoznačnost jmen atributů
- předpoklad jednoznačnosti vztahů – pro všechny množiny atributů X existuje nejvýše jeden typ vztahu

Pro návrh se používá tvrzení o bezztrátové dekompozici:

Mějme schéma $R(A,B,C)$, kde A,B,C jsou disjunktní množiny atributů a funkční závislost $B \Rightarrow C$. Rozložíme-li R na schémata $R1(B,C)$ a $R2(A,B)$, je takto provedená dekompozice bezztrátová. Naopak, je-li dekompozice $R1(B,C)$ a $R2(A,B)$ bezztrátová, musí platit buď $A \Rightarrow B$ nebo $B \Rightarrow A$.

Návrh relačního schématu ve 3. NF dekompozicí z univerzální relace

Předpokladem je

- univerzální relace
 - jednoznačnost jmen atributů
 - jméno atributu hraje jen jednu roli
- jednoznačnost vztahů

Postup:

- schémata se rozkládají binárně dle tvrzení o bezztrátové dekompozici
 - bezztrátová dekompozice má vlastnost bezztrátového spojení - po spojení získáme původní data, žádná informace se neztratí
 - je-li $R(A,B,C)$ a $A \twoheadrightarrow B$ a $B \twoheadrightarrow C$, pak bezztrátová dekompozice je $R1(B,C)$ a $R2(A,B)$
- strategie: “nalomit” tranzitivitu, dekomponovat podle funkční závislosti, která způsobuje, že schéma není v 3NF
- každé nové schéma se testuje na 3NF

Výsledek (kromě 3NF):

- je zachována bezztrátovost
- obecně nejsou pokryty závislosti

Návrh relačního schématu syntézou dle Bernsteina

Postup:

1. Určit klíč a vytvořit minimální pokrytí
2. Závislosti se roztrídí do skupiny se stejnou levou stranou
3. Sloučení schémat s ekvivalentními klíči
4. Atributy, které se nevyskytují v žádné z funkčních závislostí F buďto umístíme do samostatné relace nebo umístíme samostatně

5. Pokud žádná z relací neobsahuje klíč, připojíme k zajištění bezztrátovosti další relační schéma obsahující klíč univerzálního schématu relace.

Př.: $R(A, B, C, D, E, G)$ $BE \rightarrow C$, $BD \rightarrow C$, $A \rightarrow D$, $BC \rightarrow A$, $AE \rightarrow G$

1. klíčem je BE
2. $R = \{BEC, BDC, AD, BCA, AEG\}$
3. Sloučení BD a BC $R = \{BEC, BDCA, AD, AEG\}$
4. Neexistují samostatné atributy a ale můžeme vypustit vazbu AD protože je obsažena v BDCA
5. Protože je klíč obsažen v HSM pak výsledkem je $R = \{BEC, BDCA, AEG\}$

Objektový model

Funkční závislosti

Je definována mezi dvěma množinami atributů v rámci jednoho schématu relace. Jde tedy o vztahy mezi daty. Mějme množiny atributů $B, C \subseteq A$. pak C závisí funkčně na B, jestliže ke každé B - hodnotě existuje nejvýše jedna C - hodnota. Zapisujeme ve tvaru „ $B \rightarrow C$ “.

-
- Stránka byla naposledy editována v 20:46, 3. 2. 2007.
-