Databázové technologie

Dotazovací systémy v relačním modelu

Petr Krajča

Katedra informatiky
Univerzita Palackého v Olomouci

Relační dotazování

- **Instance databáze** je konečná množina relační proměnných, jejich aktualních hodnot a integritních omezení. (Zatím nepotřebujeme.)
- **Dotaz** je částečně rekurzivní funkce z množiny všech instancí databáze do množiny všech relací (nad relačními schématy).

Poznámky

- typický přístup: dotaz je popsán v určitém *jazyku* a je dána jeho *intepretace*
- dotazovací jazyk je **doménově nezávislý**, pokud výsledky dotazů nezávisí na typech (doménách), ale pouze na hodnotách relačních proměnných

Základní dotazovací systémy

1. **relační algebra**
   - specifikuje množinu **operací s relacemi**
   - dotaz = výraz skládájící se ze složených relačních operací
   - interpretace dotazu = postupné vyhodnocování operací
   - elementární operace s relacemi (SŘBD může dobře optimalizovat)
2. **relační kalkuly**
   - několi typů: **doménový relační kalkul**, **n-ticový relační kalkul**
   - dotaz = formule predikátové logiky (s volnými proměnnými)
   - interpretace dotazů = ohodnocení formulí ve struktuře (instanci databáze)
   - ryze deklarativní, vychází z něj řada jazyků (QUEL, částečně SQL)

Poznámka:

- databáze při zpracování dotazů vytváří **prováděcí plány**, které jsou blízké operacím relační algebry

Základní pojmy RM: Atribut, Doména

Atribut $=$ symbolické jméno. Množinu všech atributů, o které předpokládáme, že je konečená a spočetná, označujeme Y.

Typ $=$ pojmenovaná (nejvýše spočetná) množina elementů (hodnot).

- RM je silně typový, každý atribut (relace, atd.) má přiřazen svůj typ.
- ekvivalentní název pro *typ* je *doména*
- někdy uvažován rozdíl
  - **doména** – množina konkrétních hodnot (sémantický pojem)
  - **typ** – symbolické jméno pro doménu (syntaktický pojem)
- **rozlišitelnost hodnot** – hodnoty stejného typu lze porovnat $=$. Hodnoty různých typů nelze porovnat!

Základní pojmy RM: Relační schéma

Relační schéma je konečná množina $R = \{\langle y_1, \tau_1 \rangle, \dots, \langle y_n, \tau_n \rangle\}$, kde $y_1, \dots, y_n$ jsou vzájemně různé atributy z Y a $\tau_1, \dots, \tau_n$ jsou typy.

Poznámky:

- formalizace pojmu záhlaví datové tabulky
- relační schéma představuje *typ* relace (*relační typ*)
- značení: $R, R_1, R_2, \dots, S_1, S_2, \dots$
- relační schéma může být prázdné

Dohoda o značení:

Pokud typ vyplývá jednoznačně z kontextu, pak

- relační schémá ztotožňujeme s konečnými podmnožinami $R \subseteq Y$
- doménu (typ) atributu $y \in Y$ značíme D_y
- pro přehlednost budeme předpokládat, že jméno atributu implicitně určuje jeho typ

Obecný kartézský součin

Mějme systém množin A_i, které jsou indexovány prvky z množiny I (tzv. indexy).

Kartézský součin množin $A_i (i \in I)$ je množina $\prod_{i \in I} A_i$ všech zobrazení

$$f : I \to \bigcup_{i \in I} A_i$$

takových, že $f(i) \in A_i$ platí pro každý index $i \in I$.

Každé zobrazení $f \in \prod_{i \in I} A_i$ se nazývá **n-tice** (angl. tuple).

Poznámky:

- pro $f \in \prod_{i \in I} A_i$ se $f(i) \in A_i$ nazývá hodnota i v n-tici f
- pokud je $I = \{1, \dots, n\}$, pak lze psát $\prod_{i=1}^{n} A_i$
- indexy z I nemají žádné pořadí

Základní pojmy RM: Relace nad relačním schématem

Mějme relační schéma $R \subseteq Y$ a nechť D_y $(y \in Y)$ označuje domény atributů $y \in R$.

Relace $\mathcal{D}$ nad relačním schématem R je libovolná konečná podmnožina $\prod_{y \in R} D_y$.

Číslo $|\mathcal{D}|$ se nazývá **velikost relace** $\mathcal{D}$.

Číslo $|R|$ se nazývá **stupeň relace** $\mathcal{D}$.

Poznámky:

- značení: $\mathcal{D}, \mathcal{D}_1, \ldots$
- $\prod_{y \in R} D_y$ – kartézský součin domén (indexy jsou atributy)
- $r \in \prod_{y \in R} D_y$ je n-tice, $r(y)$ je prvek z D_y
- n-tice reprezentuje „řádek" v tabulce odpovídající $\mathcal{D}$
- alternativní pohled $\mathcal{D} : \prod_{y \in R} D_y \to \{0, 1\}$
- tabulka je v 1NF p.k. reprezentuje relaci na relačním schématu
- nulární relace (stupeň 0), unární relace (stupeň 1), binární relace (stupeň 2), ...

Příklady relačních operací

Pro dvě relace $\mathcal{D}_1$, $\mathcal{D}_2$ na relačním schématu $R \subseteq Y$ zavádíme

$$\begin{aligned}
\mathcal{D}_1 \cup \mathcal{D}_2 &= \{r \in \prod_{y \in R} D_y \mid r \in \mathcal{D}_1 \text{ nebo } r \in \mathcal{D}_2\} \\
\mathcal{D}_1 \cap \mathcal{D}_2 &= \{r \in \prod_{y \in R} D_y \mid r \in \mathcal{D}_1 \text{ a } r \in \mathcal{D}_2\}
\end{aligned}$$

Mějme relaci $\mathcal{D}$ na schématu R a nechť θ je výraz typu ,,pravdivostní hodnota", který může obsahovat jména atributů z R. Řekneme, že $r \in \mathcal{D}$ **splňuje** (podmínku danou výrazem) θ, pokud má θ hodnotu "pravda" za předpokladu, že jsme nahradili jména atributu v θ jejich hodnotami z r.

$$\sigma_\theta(\mathcal{D}) = \{r \in \mathcal{D} \mid r \text{ splňuje } \theta\}$$

Relace $\sigma_\theta(\mathcal{D})$ se nazývá **restrikce** $\mathcal{D}$ splňující θ.

Nticový relační kalkulus

Výrazy nticového rel. kalkulu

$$\mathcal{D} = \{x(R) \mid f(x)\}$$

- R relační schéma
- x – nticová proměnná (hodnotou je ntice)
- $f(x)$ – predikát (funkce vracející pravdivostní hodnotu pro ntici x)
- $\mathcal{D}$ – relace skládající se ze všech ntic $t(R)$, pro které je hodnota $f(t)$ pravda

Poznámka

- $x(R)$ – podle kontextu označuje schéma dané ntice, případně její zúžení; pro lepší čitelnost budeme vynechávat množinové závorky

Příklad

Relace $parts$ s relačním schématem $\{\texttt{partno}, \texttt{subpartof}, \texttt{partname}\}$.

$$\{x(\texttt{partno}, \texttt{partname}) \mid parts(x) \text{ and } x(\texttt{subpartof}) = 211\}$$

Formule nticového kalkulu (1/4)

- základem formulí nticového kalkulu jsou atomy

Atom je

1. každá nticová proměnná
2. $\mathcal{D}(x)$ s významem $x \in \mathcal{D}$, kde $\mathcal{D}$ je relační proměnná (resp. odpovídající relace)
3. $t(y_1)\ \theta\ u(y_2)$, kde t a u jsou nticové proměnné, y_1, y_2 atributy a θ je operátor porovnání
4. $c_1\ \theta\ t(y_2)$, $t(y_1)\ \theta\ c_2$, kde t je nticová proměnná, y_1, y_2 atributy, θ je operátor porovnání a hodnoty $c_1 \in D_{y_1}, c_2 \in D_{y_2}$ jsou konstanty

Příklady

- $parts(x)$
- $x(\texttt{partno}) = y(\texttt{partno})$
- $x(\texttt{amount}) > 200$

Formule nticového kalkulu (2/4)

- formule jsou definovány rekurzivně podobně jako formule predikátové logiky
- s využitím spojek
  - $\neg$ (negace)
  - $\vee$ (disjunkce)
  - $\wedge$ (konjunkce)
  - $\exists$ (existeční kvantifikátor)
  - $\forall$ (univerzální kvantifikátor)
- závorek

Formule nticového kalkulu (3/4)

Nechť f a g jsou formule, x je nticová proměnná a $R \subseteq Y$.

Formule nticového kalkulu je

1. každý atom
2. $\neg f$ je formule; $\neg f$ je pravdivá, p.k. f je nepravdivá
3. $f \vee g$ a $f \wedge g$ jsou formule; $f \vee g$ pravdivá, p.k. f nebo g je pravdivá, a $f \wedge g$ je pravdivá, pokud f a g jsou obě pravdivé
4. $\exists x(R)f$ je formule, která je pravdivá, p.k. existuje ntice t nad R taková, že nahradíme-li proměnnou x nticí t, bude f pravdivá
5. $\forall x(R)f$ je formule, která je pravdivá, p.k. pro všechny ntice t nad R platí, že nahradíme-li proměnnou x nticí t, bude f pravdivá
6. (f) je formule

Formule nticového kalkulu (4/4)

Priorita operátorů

- $\exists$ a $\forall$ mají nejvyšší a shodnou prioritu
- $\neg$, $\wedge$, $\vee$ mají prioritu nižší a v tomto pořadí

Příklady

Pro relace:

- $parts$ s relačním schématem $R_{parts} = \{\texttt{partno}, \texttt{subpartof}, \texttt{partname}\}$
- $usedon$ s relačním schématem $R_{usedon} = \{\texttt{partno}, \texttt{ptype}, \texttt{nused}\}$
- $instock$ s relačním schématem $R_{instock} = \{\texttt{partno}, \texttt{location}, \texttt{amount}\}$

- $\neg x(\texttt{amount}) \leq 10$
- $(x(\texttt{partno}) = y(\texttt{partno})) \vee \neg x(\texttt{amount}) \leq 10$
- $\exists x(\texttt{partno}, \texttt{location}, \texttt{amount})(instock(x) \wedge x(\texttt{partno}) = y(\texttt{partno}))$

Volné a vázané proměnné

- ne všechny dobře zkonstruované formule jsou validní formule nticového rel. kalkulu
- příklad $usedon(x) \vee x(\texttt{location}) = \text{"}\texttt{NYC}\text{"}$
- nekompatibilita typu proměnných

Volné a vázáné proměnné

- analogie lexikálního rozsahu platnosti (vázané proměnné $\approx$ lokální proměnné)
- kvantifikátory obstarávají vazbu proměnné a určují její typ

Určení vlastností formulí a proměnných

- nutné k určení, které formule rel. kalkulu jsou validní
- $type(x, f)$ – typ proměnné x ve formuli f
- $men(x, f)$ – mention set, množina atributů, které se vyskytly s x ve formuli f
- $type$ a men jsou definovány pouze pro volné proměnné
- využívá se rekurzivní struktura formulí

Určení vlastností formulí a proměnných (1/2)

atomické formule

Nechť f je formule ve tvaru:

1. $\mathcal{D}(x)$, pak x je volná proměnná v f a $type(x, f) = men(x, f) = R$, kde R je relační schéma $\mathcal{D}$
2. $t(y_1)\ \theta\ u(y_2)$, pak t a u jsou volné proměnné v f, $type(t, f) = type(u, v) = undefined$; $men(t, f) = \{y_1\}$, $men(u, f) = \{y_2\}$
3. $c\ \theta\ t(y)$, $t(y)\ \theta\ c$, pak t je volná proměnná, $type(t, f) = undefined$, $men(t, f) = \{y\}$

- všechny atomické formule jsou validní, pokud jsou dobře použity operátory ve vztahu k doménám

Určení vlastností formulí a proměnných (2/2)

složené formule

Nechť g a h jsou validní formule, pak formule f

1. $f = \neg g$ je validní a všechny výskyty proměnných jsou volné nebo vázané stejně jako v g, pro každou volnou proměnnou x v f platí $type(x, f) = type(x, g)$ a $men(x, f) = men(x, g)$.
2. $f = g \wedge h$ nebo $f = g \vee h$, všechny výskyty proměnných jsou volné nebo vázané stejně jako v g nebo h. Pokud je x volná pak:
   - pokud je definován typ právě pro jednu podformuli, např. $type(x, g)$ a x je volná v h, pak musí platit $type(x, g) \supseteq men(x, h)$, aby byla formule validní; $type(x, f) = type(x, g)$,
   - jinak musí platit $type(x, g) = type(x, h)$, aby formule byla validní a platí $type(x, f) = type(x, g) = type(x, h)$.
   - $men(x, f) = men(x, g) \cup men(x, h)$

3. $f = \exists x(R)g$:
   - x musí být volná proměnná v g,
   - $type(x, g)$ musí být R, pokud je definován
   - $R \supseteq men(x, g)$,
   - všechny výskyty x v f jsou vázané,
   - $type(x, f) = men(x, f) = undefined$, protože x není volná proměnná v f,
   - výskyt každé jiné proměnné $z \neq x$ v f je volný nebo vázaný tak jako v g; $type(z, f) = type(z, g)$ a $men(z, f) = men(z, g)$.
4. $f = \forall x(R)g$: dtto
5. $f = (g)$ je validní formule a platí pro ni vše co pro g

Příklady

- $\forall x(R_{instock})(\neg instock(x) \vee x(\texttt{amount}) < 100)$, validní formule, všechny proměnné jsou vázané
- $\forall x(R_{usedon})(x(\texttt{amount}) = y(\texttt{amount}))$, validní formule, proměnná x je vazaná, y je volná

Výrazy nticového kalkulu a jejich interpretace

$$\{x(R) \mid f(x)\}$$

- f – validní formule nticového relačního kalkulu
- x – jediná volná nticová proměnná v f
- R – $R \subseteq Y$
- pokud je $type(x, f) \neq undefined$, pak $type(x, f) = R$, jinak $R \supseteq men(x, f)$

Interpretace

- hodnotu výrazu získáme nahrazením (substitucí) nticových proměnných nticemi

Substituce proměnných

Nechť $f(x)$ je validní formule nticového rel. kalkulu s volnou proměnnou x, pak substituce t za x v f je označna $f(t/x)$ a představuje formuli, kde každý atom obsahující výskyt volné proměnné x byl nahrazen t následovně.

- $\mathcal{D}(x)$ je nahrazeno $true$, pokud $t \in \mathcal{D}$, jinak $false$.
- $x(y_1)\ \theta\ u(y_2)$ je nahrazeno $c\ \theta\ u(y_2)$, kde c je konstanta daná jako $t(y_1)$
- analogicky pro $u(y_1)\ \theta\ x(y_2)$
- $x(y_1)\ \theta\ x(y_2)$ je nahrazeno pravdivostní hodnotou výrazu $c_1\ \theta\ c_2$, kde $c_1 = t(y_1)$ a $c_2 = t(y_2)$
- $x(y)\ \theta\ c$ a $c\ \theta\ x(y)$ analogicky.

Příklad

- $f = (instock(x) \vee (x(\texttt{location}) = \text{"}\texttt{NYC}\text{"}) \wedge \forall y(R_{instock})(x(\texttt{partno}) = y(\texttt{partno}) \wedge y(\texttt{amount}) > 20))$
- $t = \{\langle \texttt{partno}, 2114\rangle, \langle \texttt{location}, NYC\rangle, \langle \texttt{amount}, 6\rangle\}$
- $f(t/x) = (true \vee true \wedge \forall y(R_{instock})(2114 = y(\texttt{partno}) \wedge y(\texttt{amount}) > 20))$

Interpretace formule

Nechť f je validní formule bez volných proměnných, která může obsahovat atomy $true$ a $false$. Interpretace formule $I(f)$ je definována rekurzivně. Pokud je f

1. $true$, pak $I(f) = true$.
   $false$, pak $I(f) = false$.
2. $\neg g$, pak nesmí g obsahovat žádné volné proměnné a $I(f) = true$, p.k. $I(g) = false$, jinak $I(f) = false$.
3. $g \vee h$ nebo $g \wedge h$, pak g a h nesmí obsahovat volné proměnné.
   Pro $g \wedge h$ platí $I(f) = true$ p.k. $I(g) = I(h) = true$, jinak $I(f) = false$.
   Pro $g \vee h$ platí $I(f) = true$ p.k. $I(g) = true$ nebo $I(h) = true$, jinak $I(f) = false$.
4. $\exists x(R)g$, x je jediná volná proměnná v g. $I(f) = true$ p.k. existuje alespoň jedno $t \in \prod_{y \in R} D_y$ takové, že $I(g(t/x)) = true$, jinak $I(f) = false$.
5. $\forall x(R)g$, x je jediná volná proměnná v g. $I(f) = true$ p.k. pro všechny $t \in \prod_{y \in R} D_y$ platí, $I(g(t/x)) = true$, jinak $I(f) = false$.
6. (g), pak $I(f) = I(g)$.

Hodnota relačního výrazu

- hodnota výrazu

$$\mathcal{D} = \{x(R) \mid f(x)\}$$

- je relace na relačním schématu R obsahující všechny ntice $t \in \prod_{y \in R} D_y$ splňující

$$I(f(t/x)) = true.$$