3. gyakorlat – emlékeztető, feladatok

Sidló Csaba

2005. szeptember 26.

1. Fogalmak

1.1. Relációs algebrai kifejezések ábrázolása fában

1.2. Relációs algebrai kifejezések optimalizációja

Feltesszük, hogy a relációs algebrai kifejezések kiértékelését a végrehajtási fa leveleitől elindulva a gyökér felé haladva hajtjuk végre úgy, hogy mindig kiértékeljük azokat a részfákat, amelynek már csak a gyökeréhez tartozó reláció-előfordulása nem ismert. Egy-egy művelet kiértékelésének költsége függ a művelet típusától, valamint az argumentumként kapott relációk méretétől, azok sorainak valamint oszlopainak számától.

A kifejezés optimalizációja során arra törekszünk, hogy az összesített költségek minél kisebbek legyenek, így például mepróbáljuk a részeredmény-relációk méretét minél hamarabb lecsökkenteni, direkt szorzat helyett természetes összekapcsolásokat alkalmazni stb.

A köveztkező lépéseket végrehajtva optimálisabban végrehajthatóvá tehetjük a kiértékelési fát (a pontok számozása egyfajta prioritási sorrendet tükröz):

1. a σ-kat megpróbáljuk minél lentebb vinni a fában, tehát minél hamarabb végrehajtani,
2. megpróbálunk kialakítani természetes összekapcsolásokat az

   $$r \bowtie s = \prod_{attr.lista} (\sigma_{egyenl\H{o}s\acute{e}g\ felt.} r \times s))$$

   definíció alapján,
3. ha van információnk az összekapcsolás-sorozatok tábláinak méretéről, akkor kialakítunk egy olyan végrehajtási sorrendet, ahol mindig a lehető legkisebb táblát kapcsoljuk hozzá a részeredményhez,
4. a $\prod$-ket megpróbáljuk minél lentebb vinni a fában, tehát minél hamarabb végrehajtani,
5. a kiválasztás és projekció sorozatokban megpróbáljuk az attribútumhalmazokat és feltételeket egyszerűsíteni, összevonni,
6. közös részfákat keresünk, próbálunk meg kialakítani a fában, mivel azok együttes kiértékelése várhatóan gyorsabb az egyenkénti feldolgozásuknál.

A pontokban leírt átalakításokhoz felhasználhatunk minden, bizonyítottan helyes átalakítást.

1.3. Relációs kalkulusok

A relációs algebra

- halmazorientált,
- algebrai eszközöket használ,
- valamint procedurális eszköz.

A relációs kalkulusok

- a matematikai logikán alapulnak,
- deklaratív nyelvet szolgáltatnak, melyben feltételekkel fogalmazzuk meg az elvárásainkat.

1.4. Relációs tartomány kalkulus (DRC)

DRC kifejezésre példa a következő:

$$\{(név, gyümölcs) | Sz(név, gyümölcs) \wedge név =' Mic')\}.$$

1.1. Definíció (term). *Termeknek nevezzük a DRC kifejezések alapvető építőelemeit, melyek lehetnek*

$$\begin{array}{rcl} x_1, ..., x_n & : & \textit{változók}, \\ c_1, ..., c_n & : & \textit{konstansok}. \end{array}$$

1.2. Definíció (predikátum szimbólum). *A kalkulus kifejezésekben a relációkat megfeleltetjük egy-egy predikátum szimbólumnak. Egy $P(A_1, ..., A_n)$ pred. szimbólumban $A_1, ..., A_n$ termek. Egy $P(c_1, ..., c_n)$ predikátum szimbólum akkor és csak akkor veszi fel az igaz értéket, ha a neki megfeleltetett P relációban szerepel $(c_1, ..., c_n)$ sor.*

1.3. Definíció (formula). *Elemi formuláknak nevezzük az*

- *$x_i \theta x_j$ elemi összehasonlításokat, ahol $\theta \in \{=, \neq, <, >, \leq, \geq\}$,*
- *valamint a $P(A_1, ..., A_n)$ predikátumokat.*

Ha F_1 és F_2 formulák, akkor az

- *$F_1 \wedge F_2$, $F_1 \vee F_2$, $\neg F_1$,*
- *$\forall x : F_1(x)$, $\exists x : F_1(x)$,*
- *(F)*

kifejezések szintén formulák. A DRC formulák tehát egymásba ágyazottak. Egy formulának tulajdonsága, hogy melyek a benne előforduló szabad változók. Ehhez definiálnunk kell, mikor szabad illetve kötött egy formula változója:

- *minden elemi formulában előforduló változó szabad előfordulású,*

- *a lokikai műveletekkel illetve zárójelezéssel származtatott formulák változóinak előfordulása nem változik,*
- *ha F-ben x szabad előfordulású, akkor a $\forall x : F(x)$ és $\exists x : F(x)$ formulákban x kötött előfordulású, valamint F minden kötött változója továbbra is kötött marad.*

1.4. Definíció (DRC kifejezés). *DRC kifejezéseknek hívjuk az*

$$\{x_1, ..., x_n | F(x_1, ..., x_n)\}$$

kifejezéseket, ahol F formula, $x_1, ..., x_n$ pedig F szabad előfordulású változói.

2. Feladatok

2.1. Relációs algebrai optimalizáció

Legyenek $Sz : Szeret(név, sör)$, $L, L_1, L_2 : Látogat(név, bár)$, $E : Elad(bár, sör)$ relációk. Alakítsuk át optimálisabbra a következő relációs algebrai kifejezéseket.

1. $\sigma_{név='Mic'}(\sigma_{sör='Borsodi'}(Sz))$
2. $\sigma_{bár='Pagony'}(L \bowtie E)$
3. $\sigma_{név='Mic'}(L_1 \setminus L_2)$
4. $\prod_{név}(\prod_{név,\ sör}(L \bowtie E))$
5. $\sigma_{név='Mic'}(\prod_{név,\ sör}(L \bowtie E))$

Legyenek adottak a következő relációsémák:

$$\begin{array}{rcl} F & : & Film(cím, év, hossz, stúdió, rendező) \\ J & : & Játszik(cím, év, színész) \\ V & : & Vetít(cím, év, mozinév, dátum, óra, helyár) \end{array}$$

Feltesszük, hogy a filmeket az év (amikor készítették), valamint a címe azonosítják egyértelműen. A filmek hossza percben adott. Milyen kérdésre válasz a kifejezés? Optimalizáld!

6. $\prod_{cím,év}(\sigma_{színész='XY'}(\prod_{cím,év,színész}(\sigma_{F.cím=J.cím \wedge F.év=J.év}(F \times J))))$
7. $\prod_{cím} \sigma_{v=1978}(\prod_{cím,év}(F) \cup \prod_{cím,év}(J))$
8. $\prod_{színész}(\sigma_{stúdió='Fox'}(\prod_{stúdió,hossz,színész}(\sigma_{F.hossz>100}(\sigma_{F.év=J.év \wedge F.cím=J.cím}(F \times J)))))$
9. $\prod_{stúdió}(\sigma_{F.cím='Birodalom' \wedge F.év=1993}(F \times (\prod_{cím,év}(\sigma_{J_1.színész=J_2.színész}(J_1 \times \prod_{színész}(\sigma_{J_2.cm='Birodalom' \wedge J_2.év=1993}(J_2)))))))$

2.2. Egyszerű DRC kifejezések

Legyen adott az Sz : $Szeret$(*név, gyümölcs*)) reláció. Írjunk fel DRC kifejezéseket a következő kérdésekhez.

1. Mely gyümölcsöket szereti Mic?
2. Kik szeretik az almát?
3. Mely gyümölcsöket nem szereti Mic?
4. Kik szeretnek minden gyümölcsöt?