Adatbázisok elmélete 5. előadás

Katona Gyula Y.
Budapesti Műszaki és Gazdaságtudományi Egyetem
Számítástudományi Tsz.
I. B. 137/b
kiskat@cs.bme.hu
http://www.cs.bme.hu/˜kiskat

2004

A relációs algebra alapműveletei (emlékeztető)

- Halmazműveletek (bármilyen halmazra mennének)
  - ★ unió: $\cup$
  - ★ különbség: $\setminus$
  - ★ szorzat: $\times$
- Relációs műveletek (ezek már kihasználják, hogy itt relációkról van szó)
  - ★ vetítés, projekció: π
  - ★ kiválasztás, szelekció: σ

Ezek mind tiszta műveletek: reláció → reláció
$\Longrightarrow$ gond nélkül egymásba ágyazhatók

Műveletek

Unió (emlékeztető)

- R, S relációk $\Longrightarrow$ $R \cup S$ = sorai vagy R vagy S sorai
  Azonos sorok csak egyszer szerepeljenek. (Gyakorlatban néha lehetnek azonos sorok.)
- csak akkor alkalmazható, ha R és S oszlopszáma egyenlő
- nem feltétlenül örököl típusokat vagy attribútum neveket
- Példa:

R	A	B
	a	a
	a	c
	b	a

S	A	C
	a	a
	a	d
	a	c
	b	b

$R \cup S$	A	$(R \cup S)_2$
	a	a
	a	c
	b	a
	a	d
	b	b

Műveletek

Különbség

- R, S relációk $\Longrightarrow$ $R \setminus S = R$ azon sorai, amelyek S-ben nem szerepelnek
- nincs kompatibilitási követelmény (Ha pl. különböző az oszlopszám, nem szerepelhetnek azonos sorok úgysem. Ekkor $R \setminus S = R$)
- Az eredmény örökli R típusait és attribútum neveit (mert $R \setminus S \subseteq R$)
- Példa:

R	A	B
	a	a
	a	c
	b	a

S	A	C
	a	a
	a	d
	a	c
	b	b

$R \setminus S$	A	B
	b	a

Műveletek

Szorzat (direkt szorzat, Descartes szorzat)

- $R(A_1, \ldots, A_k), S(B_1, \ldots, B_l)$ k ill. l attribútumos relációk $\Longrightarrow R \times S =$ egy $k + l$ attribútumos reláció, R minden sora mögé odatesszük S minden sorát, minden lehetséges módon.
  Ha R-nek n sora van S-nek m sora $\Longrightarrow R \times S$-nek nm sora van
- nincs kompatibilitási követelmény
- Az eredmény lényegében örökli R és S típusait és attribútum neveit, esetleg át kell nevezni.

- Példa:

R	A	B
	a	a
	a	c
	b	a

S	A	C
	a	a
	a	d
	a	c
	b	b

$R \times S$	A	B	A'	C
	a	a	a	a
	a	a	a	d
	a	a	a	c
	a	a	b	b
	a	c	a	a
	a	c	a	d
	a	c	a	c
	a	c	b	b
	b	a	a	a
	b	a	a	d
	b	a	a	c
	b	a	b	b

Az unió és különbség könnyű művelet, a szorzat nehezebb. Vigyázni kell mennyit használjuk.

Műveletek

Vetítés

- $R(A_1, \ldots, A_l)$ alakú reláció $\Longrightarrow \pi_{A_{i_1},\ldots,A_{i_n}}(R)$
  R vetítése $A_{i_1}, \ldots, A_{i_n}$-re (fontos a sorrend) $\Longrightarrow$
  Veszem az oszlopokat ebben a sorrendben, a többit eldobom és a többszörös sorokat is eldobom.
  Egy oszlop akár többször is szerepelhet. $\Longrightarrow$ átnevezés
- nincs kompatibilitási követelmény (persze amire vetítünk az R-nek attribútuma kell, hogy legyen)
- Az eredmény örökli R típusait és attribútum neveit
- Példa:

R	A	B	C
	a	b	2
	a	c	3
	b	c	4

$\pi_A(R)$	A
	a
	b

$\pi_{C,B,B}(R)$	C	B	B
	2	b	b
	3	c	c
	4	c	c

Műveletek

Kiválasztás, szelekció

- R egy reláció $\Longrightarrow \sigma_F(R) =$ a reláció azon sorai, amelyekre az F formula teljesül.
- Teljesülni fog, hogy $\sigma_F(R) \subseteq R$
- Nincs megszorítás, csak hogy F értelmes legyen.
- Az eredmény örökli R típusait és attribútum neveit
- Példa:

R	A	B	C
	a	b	2
	a	c	3
	b	c	4

$\sigma_{A \neq B \wedge C>2}(R)$	A	B	C
	a	c	3
	b	c	4

Az F formula:

Atomok: $A\ \theta\ B$, $A\ \theta\ c$, $c\ \theta\ B$,
ahol A, B attribútumok, c érték (konstans), $\theta \in \{<, >, =, \boxed{\leq, \geq, \neq}\}$

Építkezés: $\wedge, \vee, \neg, (,)$ Kvantorok, nincsenek!

- Példa:

DOLGOZÓ(NÉV,CÍM,FIZETÉS)

$$\sigma_{\text{CÍM='BP., Várna u.'} \wedge \text{FIZETÉS>'50000'}}(\text{DOLGOZÓ})$$

Relációs algebra

Definíció. *Alapreláció: A bevezetés, tervezés során definiált tábla, ami meg van adva.*
A relációs algebra relációi: amik kifejezhetők az alaprelációkból $\cup, \setminus, \times, \pi, \sigma$ *segítségével.*
Származtatott reláció: nem alapreláció, de kifejezhető.
Definíció. *Egy lekérdező nyelv (igazi vagy modell) relációsan teljes, ha benne megvalósíthatók a relációs algebra alapműveletei:* $\cup, \setminus, \times, \pi, \sigma$

Ez fontos követelmény, általában tudja is mindegyik.
Inkább az a baj, hogy néha túl sokat tudnak, de nincs hatékony implementáció.

Származtatott műveletek

Hasznosak, de mivel kifejezhetők az 5 alapművelettel, ezért lényegében csak rövidítések.

Metszet

- R, S relációk $\Longrightarrow R \cap S = R \setminus (R \setminus S)$ azok a sorok, amelyek mindkettőben benne vannak.
- nincs kompatibilitási követelmény $\Longleftarrow \setminus$ tulajdonságából
- Az eredmény örökli R típusait és attribútum neveit
  $\Longleftarrow \setminus$ tulajdonságából
- Példa:

R	A	B
	a	a
	a	c
	b	a

S	A	C
	a	a
	a	d
	a	c
	b	b

R ∩ S	A	B∩C
	a	a
	a	c

Származtatott műveletek

Természetes illesztés (natural join)

- $R(A_1, \ldots, A_k, B_1, \ldots, B_r), S(A_1, \ldots, A_k, C_1, \ldots, C_s)$ relációk
  $\Longrightarrow R \bowtie S =$
  - ★ Vegyük $R \times S$-t
  - ★ Vesszük azokat a sorokat, ahol $R.A_1 = S.A_1, \ldots, R.A_k = S.A_k$, a többit kidobjuk.
  - ★ $\forall A_i$-ból az egyik példányt eldobjuk, azaz vetítünk $R.A_1, \ldots, R.A_k, R.B_1, \ldots, R.B_r, S.C_1, \ldots, S.C_s$-re
  - ★ Azonos sorokat kidobjuk.
- $$R \bowtie S = \pi_{R.A_1,\ldots}(\sigma_{R.A_1=S.A_1,\ldots}(R \times S))$$

  $R \bowtie S$-nek $k + r + s$ oszlopa lesz.
  Ha nincs közös attribútum. $\Longrightarrow R \bowtie S = R \times S$.
- nincs kompatibilitási követelmény
- Az eredmény örökli R és S típusait és attribútum neveit
- Gyakorlatban ennél hatékonyabban számítjuk ki.
- Az oszlopok sorrendje nem definiált, de általában: R oszlopai, aztán S saját oszlopai.

- Példa:

R	A	B	C
	a	b	2
	a	c	3
	b	a	4

S	D	C
	a	2
	b	3
	x	2

$R \bowtie S$	A	B	C	D
	a	b	2	a
	a	b	2	x
	a	c	3	b

Miért „természetes”?

Példa: TERMELŐ(TNÉV,TERMÉK,ÁR,CÍM)

- TNÉV → CÍM
- TNÉV, TERMÉK → ÁR

Gond: TERMELŐ címét minden terméknél tároljuk
$\Longrightarrow$ redundancia + veszélyek : cím mindig kell, minden módosításhoz; könnyen sérülhet a fent megadott függés, ha elírom a címet; akkor is kell a cím, ha csak új árut akarok felvenni)

Megoldás: Inkább tároljuk két táblában:

$R = \pi_{\text{TNÉV, CÍM}}$(TERMELŐ) és
$S = \pi_{\text{TNÉV, TERMÉK, ÁR}}$(TERMELŐ)

$\Longrightarrow$ TERMELŐ=$R \bowtie S$ (ha kell egyben a tábla, vissza lehet állítani)

Jó-e bármilyen felbontás?

$R' = \pi_{\text{TNÉV, CÍM, ÁR}}$(TERMELŐ) és
$S' = \pi_{\text{TNÉV, TERMÉK}}$(TERMELŐ)

$\Longrightarrow$ minden terméknek ugyanannyi lesz az ára (sok ára lesz)
$\Longrightarrow$ TERMELŐ$\subsetneq R' \bowtie S'$

Az lesz majd a kérdés, hogy mik lesznek a jó felbontások?

Származtatott műveletek

Bal (jobb) félillesztés

- $R(A_1, \dots, A_k, B_1, \dots, B_r), S(A_1, \dots, A_k, C_1, \dots, C_s)$ relációk
  $\Longrightarrow$ $R \ltimes S =$ R azon sorai, amelyhez vannak passzoló sorok S-ben
  $R \ltimes S = \pi_R(R \bowtie S)$
- $R \ltimes S \subseteq R$
- $R \rtimes S =$ ugyanez jobbról
-

R	A	B	C
	a	b	2
	a	c	3
	b	a	4

S	D	C
	a	2
	b	3
	x	2

$R \ltimes S$	A	B	C
	a	b	2
	a	c	3

$R \rtimes S$	D	C
	a	2
	b	3
	x	2

Származtatott műveletek

θ-illesztés

- R, S relációk
  $\Longrightarrow R \underset{R.A_i \theta S.B_j}{\bowtie} S = R \times S$ azon sorai, amelyben az adott oszlopok θ relációban vannak
  $R \underset{R.A_i \theta S.B_j}{\bowtie} S = \sigma_{R.A_i \theta S.B_j}(R \times S)$
- Példa:

R	A	B	C
	a	b	2
	a	c	3
	b	a	4

S	D	E
	a	2
	b	3
	x	2

$R \underset{C \leq E}{\bowtie} S$	A	B	C	D	E
	a	b	2	a	2
	a	b	2	b	3
	a	b	2	x	2
	a	c	3	b	3