# Relációs algebra

Ullman-Widom: Adatbázisrendszerek Alapvetés

Második, átdolgozott kiadás, Panem, 2009

2.4. Egy algebrai lekérdező nyelv

2.5. Relációkra vonatkozó megszorítások

# Mit nevezünk algebrának?

- Egy algebra általában műveleteket és atomi operandusokat tartalmaz.
- Az algebra lehetővé teszi kifejezések megfogalmazását az atomi operandusokon és az algebrai kifejezéseken végzett műveletek alkalmazásával kapott relációkon.
- Fontos tehát, hogy minden művelet végeredménye reláció, amelyen további műveletek adhatók meg.
- A relációs algebra atomi operandusai a következők:
  - a relációkhoz tartozó változók,
  - konstansok, amelyek véges relációt fejeznek ki.

# Relációs algebra (műveletek) I.

- Projekció (vetítés). Adott relációt vetít le az alsó indexben szereplő attribútumokra.
- $\mathbf{X \subseteq \{A1,\ldots,An\}}$
- $\Pi_X(r)$ sémája X
- $\Pi_X(r) := \{ t \mid \text{van olyan } t' \in r, \text{ melyre } t'[X] = t \}$
- Példa: $\Pi_{A,\,B}(r)$

| A | B | C |
|---|---|---|
| a | b | c |
| c | d | e |
| g | a | d |

$\Rightarrow$

| A | B |
|---|---|
| a | b |
| c | d |
| g | a |

# Relációs algebra (műveletek) II.

- Szelekció (kiválasztás). Kiválasztja az argumentumban szereplő reláció azon sorait, amelyek eleget tesznek az alsó indexben szereplő feltételnek.
- $\sigma_F(r)$ és r sémája megegyezik
- $\sigma_F(r) := \{ t \mid t \in r \text{ és } F(t) = \text{IGAZ} \}$
- $R(A_1, \ldots, A_n)$ séma feletti r reláció esetén a $\sigma_F$ kiválasztás F feltétele a következőképpen épül fel:
  - atomi feltétel: $A_i \ \theta \ A_j$, $A_i \ \theta$ c, ahol c konstans, $\theta \in \{=, <, >\}$,
  - ha $B_1$, $B_2$ feltételek, akkor $\neg B_1$, $B_1 \wedge B_2$, $B_1 \vee B_2$ is feltételek.
- Példa: $\sigma_{A=a \vee C=d}(r)$

| A | B | C |
|---|---|---|
| a | b | c |
| c | d | e |
| g | a | d |

→

| A | B | C |
|---|---|---|
| a | b | c |
| g | a | d |

# Relációs adatmodell (műveletek) III.

- Mivel sorok halmazáról van szó, így értelmezhetők a szokásos halmazműveletek: az unió, a metszet és a különbség.
- r, s és $r \cup s$ azonos sémájú
- $\mathbf{r} \cup \mathbf{s}$ := {t | t∈ vagy t∈s} és $\mathbf{r}$ - $\mathbf{s}$ := { t | t ∈ r és t ∉ s}
- Az alapműveletekhez az unió és különbség tartozik, metszet műveletet származtatjuk $\mathbf{r \cap s = r - (r - s)}$
- Az alábbi példa: r – s

r

| A | B | C |
|---|---|---|
| a | b | c |
| c | d | e |
| g | a | d |

s

| A | B | C |
|---|---|---|
| a | b | c |
| c | d | e |
| g | d | f |

| A | B | C |
|---|---|---|
| g | a | d |

# Relációs algebra (műveletek) IV.

- A Descartes-szorzat is értelmezhető (ezt nem tekintjük alapműveletnek, csak a természetes összekapcsolást). Itt természetesen nem fontos az attribútumok egyenlősége. A két vagy több reláció azonos nevű attribútumait azonban meg kell különböztetni egymástól.
- $r \times s := \{ t \mid t[R] \in r \text{ és } t[S] \in s \}$
- Példa: $r \times s$.

| A | B | C |
|---|---|---|
| a | b | c |
| c | d | e |
| g | a | d |

| B | D |
|---|---|
| b | r |
| q | s |

→

| A | R.B | C | S.B | D |
|---|---|---|---|---|
| a | b | c | b | r |
| a | b | c | q | s |
| c | d | e | b | r |
| c | d | e | q | s |
| g | a | d | b | r |
| g | a | d | q | s |

# Relációs algebra (műveletek) V.
# Természetes összekapcsolás

- Természetes összekapcsolás: $R(A_1,\ldots,A_n)$, $S(B_1,\ldots,B_m)$ sémájú r és s táblák esetén r |X| s azon sorpárokat tartalmazza r-ből illetve s-ből, amelyek R és S azonos attribútumain megegyeznek.
- r, s sémái R(A1,…,An,B1,…,Bk), illetve S(B1,…,Bk,C1,…,Cm)
- r |X| s =

$$\rho_{P(A1,\ldots,An,B1,\ldots,Bk,C1,\ldots,Cm)}\Pi_{A1,\ldots,An,R.B1,\ldots,R.Bk,C1,\ldots,Cm}\sigma_{R.B1=S.B1\wedge\ldots\wedge R.Bk=S.Bk}(r\times s)$$

| A | B |
|---|---|
| 0 | 0 |
| 2 | 1 |
| 1 | 2 |

$\bowtie$

| B | C |
|---|---|
| 0 | 0 |
| 0 | 2 |
| 1 | 3 |
| 4 | 3 |

=

| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 2 |
| 2 | 1 | 3 |

# Miért olyan gyakori?

Felszolgál

| kocsma | sör |
|---|---|
| Makk 7-es | Dreher |
| Lórúgás | Kozel |
| Lórúgás | Gösser |

|X|

Látogat

| név | kocsma |
|---|---|
| Péter | Makk 7-es |
| Feri | Lórúgás |

⇓

| kocsma | sör | név |
|---|---|---|
| Makk 7-es | Dreher | Péter |
| Lórúgás | Kozel | Feri |
| Lórúgás | Gösser | Feri |

A természetes összekapcsolás kifejezhető a többi alapművelettel:

$R \mid X \mid S \equiv \pi_L(\sigma_C (R \times S))$,

itt: C a közös attribútumok egyenlőségét írja elő, L pedig csak egyszer veszi fel a közös attribútumokat.

# Relációs algebra (műveletek) VI.

- Egyes esetekben szükség lehet egy adott reláció attribútumainak **átnevezésére**. A $\rho_{S(C, D, E)}(r)$ az R(A, B, C) séma feletti r reláció helyett veszi az S relációt, melynek sorai megegyeznek R soraival, az attribútumai pedig rendre C, D, E.
- Ha az attribútumokat nem szeretnénk átnevezni, csak a relációt, ezt $\rho_S(R)$-rel jelöljük.

# További (származtatott) műveletek Théta-összekapcsolás I.

- A gyakorlatban szinte kizárólag valamilyen összekapcsolásra visszavezethető műveletet használnak abban az esetben, amikor a lekérdezés megválaszolásához több táblából kell kigyűjteni az adatokat.
- Théta-összekapcsolás: $R(A_1,\ldots,A_n)$, $S(B_1,\ldots,B_m)$ sémájú táblák esetén:

  $R\ |X|_F S = \sigma_F(R \times S)$ teljesül, itt F
  - elemi feltétel $Ai\ \Theta\ Bj$, $Ai\ \Theta\ c$, ahol $\Theta \in \{=,<,>\}$ és c konstans,
  - vagy összetett feltétel, azaz: ha $B_1$, $B_2$ feltétel, akkor $\neg B_1$, $B_1 \wedge B_2$, $B_1 \vee B_2$ is feltétel.

# Théta-összekapcsolás II.

| A | B | C |
|---|---|---|
| a | b | c |
| d | d | g |
| e | f | r |

$|X|_{A = D}$

| D | E |
|---|---|
| b | e |
| d | r |

| A | B | C | D | E |
|---|---|---|---|---|
| d | d | g | d | r |

- Egyen-összekapcsolás (equi join): ha a théta-összekapcsolásban a Θ helyén = szerepel.

# Félig-összekapcsolás

- r, s sémái R(A1,…,An,B1,…,Bk), illetve S(B1,…,Bk,C1,…,Cm)
- $r \ltimes s = \rho_{P(A1,\dots,An,B1,\dots,Bk,)}\Pi_{A1,\dots,An,R.B1,\dots,R.Bk}\,(r|\times|s)$
- Az első relációban mely sorokhoz létezik kapcsolható sor a második táblából

| A | B |
|---|---|
| 0 | 0 |
| 2 | 1 |
| 1 | 2 |
| | |

$\ltimes$

| B | C |
|---|---|
| 0 | 0 |
| 0 | 2 |
| 1 | 3 |
| 4 | 3 |

=

| A | B |
|---|---|
| 0 | 0 |
| 2 | 1 |

# Külső összekapcsolás

- Nem relációs algebrai művelet, mert kilép a modellből
  - r, s sémái R(A1,…,An,B1,…,Bk), illetve S(B1,…,Bk,C1,…,Cm)
  - r $\overset{o}{\bowtie}$ s = r $\bowtie$ s relációt kiegészítjük az r és s soraival, a hiányzó helyekre NULL értéket írva

| A | B |
|---|---|
| 0 | 0 |
| 2 | 1 |
| 1 | 2 |

$\overset{o}{\bowtie}$

| B | C |
|---|---|
| 0 | 0 |
| 0 | 2 |
| 1 | 3 |
| 4 | 3 |

=

| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 2 |
| 2 | 1 | 3 |
| 1 | 2 | NULL |
| NULL | 4 | 3 |

# Összekapcsolások

- Ha r, s **sémái megegyeznek**, akkor $\mathbf{r|\times|s = r \cap s}$.
- Ha r, s sémáiban **nincs közös attribútum**, akkor $\mathbf{r|\times|s = r\times s}$.
- Ha $r = \emptyset$, akkor $r\times s = \emptyset$ és $r|\times|s = \emptyset$.
- A külső összekapcsolás lehet bal oldali, ha csak r sorait vesszük hozzá a természetes összekapcsolás-hoz: $\mathbf{r|\overset{o}{\times}|_B s}$. Hasonlóan értelmezhetjük a jobb oldali összekapcsolást is $\mathbf{r|\overset{o}{\times}|_J s}$.

## Osztás, hányados

- Maradékos osztás: $7 \div 3 = 2$, mert 2 a legnagyobb egész, amelyre még $2 * 3 \leq 7$.
- Relációk szorzata esetén $\leq$ helyett tartalmazás.
- r és s sémája **R**(**A1,…,An**,**B1,…,Bm**), illetve **S**(B1,…,Bm), **r ÷ s** sémája R(**A1,…,An**)
- **r ÷ s** a legnagyobb (**legtöbb sort tartalmazó**) reláció, amelyre $(r \div s) \times s \subseteq r$.
- Kifejezhető relációs algebrában:
- $\Pi_{A1,\ldots,An}(r) - \Pi_{A1,\ldots,An}(\Pi_{A1,\ldots,An}(r) \times s - r)$
- Lehetséges értékekből kivonjuk a rossz értékeket.
- $(p \times r) \div r = p$

# Osztás, hányados

- Ki szereti legalább azokat, mint Micimackó?

| KI | MIT |
|---|---|
| Füles | málna |
| Füles | méz |
| Füles | alma |
| Micimackó | málna |
| Micimackó | méz |
| Kanga | málna |
| Kanga | körte |
| Nyuszi | lekvár |

÷

| MIT |
|---|
| málna |
| méz |

=

| KI |
|---|
| Füles |
| Micimackó |

$$\mathbf{szeret} \div \Pi_{\mathbf{MIT}}(\sigma_{\mathbf{KI='Micimackó'}}(\mathbf{szeret}))$$

r(a,b)÷s(b) hányados kifejezése:

$$\mathbf{r(a,b)\div s(b)=\Pi_a(r)-\Pi_a(\Pi_a(r)\times s-r)}$$

# Relációs algebra, mint relációs lekérdező nyelv

Relációs algebrai kifejezések:

- elemi kifejezések

  -- konstans reláció (konkrét tábla)

  -- relációnév R (értéke R előfordulása)

- ha $E_1$, $E_2$ kifejezések, akkor a következők is kifejezések

  -- $E_1 \cup E_2$ unió

  -- $E_1 - E_2$ különbség

  -- $\Pi_{A,\, B}(E)$ vetítés

  -- $\sigma_F(E)$ kiválasztás

  -- $E_1 \,|X|\, E_2$ természetes összekapcsolás

  -- $\rho_{S(A1,\, \ldots,\, An)}(E)$ átnevezés

- Ezek és csak ezek a kifejezések, amit így meg tudunk adni.

# Relációs algebrai kifejezés kifejezésfája

- A táblák legyenek:
  Film (cím, év, hossz)
  Szerepel (filmcím, év, színésznév)
  Színész (név, kor, város)

- Adjuk meg relációs algebrai kifejezéssel, hogy a nem budapesti, 40 évesnél idősebb színészek milyen hosszú filmekben játszottak 1998-ban! Rajzoljuk fel a relációs algebrai kifejezésnek megfelelő kifejezésfát (a fa kiértékelése alulról felfelé történik)

# Relációkra vonatkozó megszorítások megadása a relációs algebra segítségével

- A megszorításokat kétféleképpen fejezhetjük ki (legyenek R és S relációs algebrai kifejezések):
  - $R = \emptyset$, azaz R-nek üresnek kell lennie,
  - $R \subseteq S$, azaz R eredményének minden sorának benne kell lennie S eredményében.
- A két megszorítás kifejezőerő szempontjából azonos:
  - $R \subseteq S$ így is kifejezhető: $R - S \subseteq \emptyset$,
  - míg $R = \emptyset$, $R \subseteq \emptyset$ alakban is írható.

# Relációkra vonatkozó megszorítások

- Hivatkozási épség megszorítás: ha egy érték megjelenik valahol egy környezetben, akkor ugyanez az érték egy másik, az előzővel összefüggő környezetben is meg kell, hogy jelenjen.
- Példa: a **SzerepelBenne**(filmCím, filmÉv, SzinészNév), **Filmek**(filmcím, év, hossz, műfaj, stúdióNév, prodAzon) táblák esetén megköveteljük, hogy a megszorítást: $\Pi_{\text{filmCím, filmÉv}}(\text{SzerepelBenne}) \subseteq \Pi_{\text{filmcím, év}}(\text{Filmek})$.
- Általában: $\Pi_A(R) \subseteq \Pi_B(S)$.
- Kulcsmegszorítás illetve további megszorítások, (például tartománymegszorítás) kifejezhetők relációs algebrában.