

Adatbázisok elmélete 7. előadás

Csima Judit
 Budapesti Műszaki és Gazdaságtudományi Egyetem
 Számítástudományi Tsz.
 I. B. 136/b
 csima@cs.bme.hu

2003. Március 5.

Származtatott műveletek

θ -illesztés

- R, S relációk
 $\implies R \bowtie_{R.A_i \theta S.B_j} S = R \times S$ azon sorai, amelyben az adott oszlopok θ relációban vannak

$$R \bowtie_{R.A_i \theta S.B_j} S = \sigma_{R.A_i \theta S.B_j}(R \times S)$$

- Példa:

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

S	D	E
a	2	
b	3	
x	2	

$R \bowtie_{C \leq E} S$	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	

2

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
 $\implies R \ltimes S = R$ azon sorai, amelyhez vannak passzolók sorok S -ben
 $R \ltimes S = \pi_R(R \times S)$
- $R \ltimes S \subseteq R$
- $R \ltimes S =$ ugyanez jobbról
- Ha R és S ugyanaz, mint múltkor:

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

$R \ltimes S$	D	E
a	2	
b	3	
x	2	

1

Példák relációs algebra alkalmazására

ÁRU(ÁRUKÓD, ÁRUNÉV, EGYSÉGÁR)
 MENNYISÉG(DÁTUM, ÁRUKÓD, DB)
 BEVÉTEL(DÁTUM, ÖSSZEG)
 BEFIZ(ÖSSZEG, BEFIZ) BEFIZ=ÖSSZEG-4000

A 2002. jan. 1. utáni napok bevételei a dátummal együtt:

$$\sigma_{\text{DÁTUM} > '2002-01-01'}(\text{BEVÉTEL})$$

A 2002. jan. 15-i befizetett összeg és bevétel:

$$\pi_{\text{ÖSSZEG, BEFIZ}}(\sigma_{\text{DÁTUM} = '2002-01-15'}(\text{BEVÉTEL} \ltimes \text{BEFIZ}))$$

$$\pi_{\text{ÖSSZEG, BEFIZ}}(\sigma_{\text{DÁTUM} = '2002-01-15'}(\text{BEVÉTEL}) \ltimes \text{BEFIZ})$$

Hány darabot adtak el 2002. jan. 15-én az A123 kódú áruból, mi a neve és az ára?

$$\pi_{\text{DB, ÁRUNÉV, EGYSÉGÁR}}(\sigma_{\text{ÁRUKÓD} = 'A123' \wedge \text{DÁTUM} = '2002-01-15'}(\text{MENNYISÉG} \ltimes \text{ÁRU}))$$

3

$$\pi_{DB, \text{ÁRUNÉV, EGYSÉGÁR}} \left(\sigma_{\text{ÁRUKÓD}='A123' \wedge \text{DÁTUM}='2002-01-15'} (\text{MENNYISÉG}) \bowtie \text{ÁRU} \right)$$

Mely nevű áruk azok, amelyekkel van azonos egységárú másik áru?

Az ÁRU reláció két sorát kell összevetni.

Átnevezés

- Technikai segítség, ha pl. két relációban ugyanolyan attribútum név van, és direkt szorzatot akarunk. Nem változtatja meg a reláció sorait, csak az attribútumok és a reláció nevét, ezért nem igazi művelet.
- $R(A_1, \dots, A_n)$ egy reláció
 $\implies \rho_S(B_1, \dots, B_n)(R) =$ sorai megegyeznek R soraival, a reláció neve S , attribútumai rendre B_1, \dots, B_n .
- Ha csak a relációt akarjuk átnevezni: $\rho_S(R)$

4

További példák

TERMÉK(GYÁRTÓ, MODELL, TÍPUS)

PC(MODELL, SEBESSÉG, MEMÓRIA, MEREVLEMEZ, CD, ÁR)

LAPTOP(MODELL, SEBESSÉG, MEMÓRIA, MEREVLEMEZ, KÉPERNYŐ, ÁR)

NYOMTATÓ(MODELL, SZÍNES, TÍPUS, ÁR)

A relációk jelentése:

TERMÉK: az adott nevű gyártó gyártja az adott modellszámú és adott típusú (PC, Laptop vagy nyomtató) terméket

PC: modellszám, sebesség megaHz-ben, memória megabájtban, merevlemez gigabájtban, a CD sebessége (pl. 4x), az ár

Laptop: mint PC-nél, plusz a képernyő mérete hüvelykben

Nyomtató: modellszám, színes-e (i/n), típusa (tintasugaras, lézer, mátrix), ára

A modellszámokról feltesszük, hogy egyediek.

6

Megoldás:

$$\text{ÁRU1} = \rho_{\text{ÁRU1}(\text{ÁRUKÓD1, ÁRUNÉV1, EGYSÉGÁR1})}(\text{ÁRU})$$

$$\text{ÁRU2} = \rho_{\text{ÁRU2}(\text{ÁRUKÓD2, ÁRUNÉV2, EGYSÉGÁR2})}(\text{ÁRU})$$

$$\text{ÁRU3} = \text{ÁRU1} \bowtie_{\text{EGYSÉGÁR1} = \text{EGYSÉGÁR2} \wedge \text{ÁRUKÓD1} \neq \text{ÁRUKÓD2}} \text{ÁRU2}$$

$$\text{ÁRU4} = \pi_{\text{ÁRUNÉV1}}(\text{ÁRU3})$$

5

- Melyek azok a PC modellek, amelynek sebessége legalább 150?

$$\pi_{\text{MODELL}} (\sigma_{\text{SEBESSÉG} \geq 150} (\text{PC}))$$

- Mely gyártók készítenek legalább egy gigás merevlemezű laptopot?

$$\pi_{\text{GYÁRTÓ}} (\text{TERMÉK} \bowtie_{\sigma_{\text{MEREVLEMEZ} \geq 1}} (\text{LAPTOP}))$$

- Adjuk meg a B gyártó által gyártott összes termék modellszámát és árát típustól függetlenül!

$$\pi_{\text{MODELL, ÁR}} (\sigma_{\text{GYÁRTÓ}='B' \wedge \text{TÍPUS}='PC'} (\text{TERMÉK}) \bowtie \text{PC}) \cup$$

$$\pi_{\text{MODELL, ÁR}} (\sigma_{\text{GYÁRTÓ}='B' \wedge \text{TÍPUS}='LAPTOP'} (\text{TERMÉK}) \bowtie \text{LAPTOP}) \cup$$

$$\pi_{\text{MODELL, ÁR}} (\sigma_{\text{GYÁRTÓ}='B' \wedge \text{TÍPUS}='NYOMTATÓ'} (\text{TERMÉK}) \bowtie \text{NYOMTATÓ})$$

- Melyek azok a gyártók, akik laptopot gyártanak, de PC-t nem?

$$\text{TERMÉK1} = \rho_{\text{TERMÉK1}} (\pi_{\text{GYÁRTÓ, TÍPUS}} (\text{TERMÉK}))$$

$$\pi_{\text{GYÁRTÓ}} (\sigma_{\text{TÍPUS}='LAPTOP'} (\text{TERMÉK1})) \setminus \pi_{\text{GYÁRTÓ}} (\sigma_{\text{TÍPUS}='PC'} (\text{TERMÉK1}))$$

7

- Melyek azok a gyártók, amelyek gyártanak legalább két, egymástól különböző, legalább 133 Mhz-en működő PC-t vagy Laptopot? (Nincs két azonos modellszám!)

$$R1 = \pi_{\text{MODELL, SEBESSÉG}}(\text{PC}) \cup \pi_{\text{MODELL, SEBESSÉG}}(\text{LAPTOP})$$

$$R2 = \pi_{\text{GYÁRTÓ, MODELL}}(\sigma_{\text{SEBESSÉG} \geq 133}(R1) \bowtie \text{TERMÉK})$$

$$R3 = \rho_{R3}(\text{GYÁRTÓ2, MODELL2})(R2)$$

$$R4 = R2 \bowtie_{\text{GYÁRTÓ} = \text{GYÁRTÓ2} \wedge \text{MODELL} \neq \text{MODELL2}} R3$$

$$R5 = \pi_{\text{GYÁRTÓ}}(R4)$$

Megjegyzés: kifejezésfával is meg lehet adni a relációs algebrai kifejezéseket, de ezeket most nem rajzolom ide le....

A NULL érték

Lehet, hogy vannak kitöltetlen mezők, ezt meg akarjuk engedni. Pl. ha csak a termelő anyjának neve hiányzik, attól még a termék ára kell. \implies NULL érték

2 alapvető értelmezés (majd SQL-nél lesz, hogy hogyan kell megmondani, hogy melyik van éppen, illetve, hogy lehet-e egyáltalán NULL valahol):

- \nexists
- \exists , de nem ismerjük.

Attól függően, hogy hogyan értelmezzük a NULL-t:

Mi legyen egy ilyen kérdéssel?:

Pl. $\pi_{\text{CÍM}=\text{'BP'}}(\text{TERMELŐ})$

Ilyenkor belevegyük-e ha a cím NULL?

További műveletek

Ezek nincsenek benne a relációs algebraiban, de fontosak, nem túl műveletigényesek.

aggregátumok: MIN, MAX, AVG, SUM, CNT (szám)

Pl. leggyorsabb gép, átlagár, hányféle printer
eredmény mindig egy szám

aggregátum csoportosítva: Bizonyos feltételek szerinti partíciókban aggregátumok.

Pl. átlagos ár tintasugaras nyomtatók között, egy gyártónak hány terméke van
 \implies eredmény egy reláció pl. (gyártó, szám) párokból.

rekurzív lezárás: (hagyományos adatkezelésben ritka, intelligensebb rendszerekben inkább)

Pl. reláció: ki főnöke kinek \implies lezárás: ki felettese kinek

reláció: melyik városból melyikbe van repülő járat \implies lezárás: átszállással el lehet-e jutni

Ezt a relációs algebra nem tudja, csak fix mélységre: pl. max 4 átszállás, (SQL nem igazán támogatja, de azért kicsit igen: SQL3, majd lesz)

Külső illesztés (outer join)

Definíció. R, S relációk $\implies R \bowtie S$ **bal külső illesztés:** $R \bowtie S$ -hez azokat az R -beli sorokat is hozzáveszük, amihez nem illeszkedik S -beli. Hiányzó helyekre NULL kerül.

Pl. SZEMÉLY(NÉV, KÓD), CÍM(K ÓD, CÍM)

SZEMÉLY \bowtie CÍM \implies akinek nincs címe nem lesz rajta

SZEMÉLY \bowtie CÍM \implies kiderül, kinek nincs meg a címe

SQL-ben van, relációs algebraival elvileg nem fejezhető ki (NULL miatt), de elkerülhető. Lényegében: $(R \bowtie S) \cup (R \setminus (R \bowtie S))$

Van **jobb külső illesztés** is: $R \ltimes S$

Teljes külső illesztés: $R \ltimes S := (R \bowtie S) \cup (R \ltimes S)$

Példa:

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

S	D	C
	a	2
	b	3
	x	2
	y	1

$R \bowtie S$	A	B	C	D
	a	b	2	a
	a	b	2	x
	a	c	3	b
	b	a	4	NULL

$R \ltimes S$	A	B	C	D
	a	b	2	a
	a	b	2	x
	a	c	3	b
	NULL	NULL	1	y

$R \bowtie \ltimes S$	A	B	C	D
	a	b	2	a
	a	b	2	x
	a	c	3	b
	b	a	4	NULL
	NULL	NULL	1	y

Multihalmazos szemantika

A relációs algebrában ugyan minden reláció halmaz, ezért nincsenek többszörös sorok, de pl. SQL-nél lesznek. A multihalmazokkal kicsit máshogy vannak a halmazműveletek:

Ha a t sor $m_R(t)$ példányban van meg R -ben és $m_S(t)$ példányban van meg S -ben, akkor

- $m_{(R \cup S)}(t) := m_R(t) + m_S(t)$ példányban lesz meg R és S úniójában
- $m_{(R \cap S)}(t) := \min\{m_R(t), m_S(t)\}$ példányban lesz meg R és S metszetében
- $m_{(R \setminus S)}(t) := \max\{m_R(t) - m_S(t), 0\}$ példányban lesz meg $R \setminus S$ -ben

Külső unió

Részben kompatibilis relációk egyesítésére:

DIÁK(NÉV, TÉMAVEZ, TSZK)

TANÁR(NÉV, TSZK, BEOSZT)

	DIÁK \cup_k TANÁR			
	NÉV	TSZK	TÉMAVEZ	BEOSZT
diák				NULL
tanár			NULL	

ODL séma átírása relációsémává

Legegyszerűbb eset \implies az osztályoknak csak attribútumai vannak, amik atomi típusúak:

```
interface Film (keys (cím, év)){
    attribute cím;
    attribute év;
    attribute hossz;
    attribute szalagFajta;
};
 $\implies$  Film(cím, év, hossz, szalagFajta)
```

A kulcs az ODL-es kulcs lesz, ha egy van csak belőle; ezt aláhúzással jelöljük. Ha több kulcs van: egyiket jelöljük, a többit írásban (de persze azok is kulcsok a relációsémában is).

Összetett típusú attribútumok: pl. rekordszerkezet OK \implies

ADATBÁZISOK ELMÉLETE 7. ELŐADÁS

```
interface Színész (key név){  
    attribute String név;  
    attribute Struct Cím  
        { string város, string utca } lakcím;  
};  
⇒ Színész(név, város, utca)
```

Kulcs: ugyanúgy, mint a nem összetett attribútumnál

De pl. halmaz csak rosszul oldható meg: halmaz minden eleméhez új sor.

```
interface Színész {  
    attribute String név;  
    attribute Set <  
        Struct Cím{ string város, string utca }  
        > lakcím;  
};  
⇒ Színész(név, város, utca)
```

ADATBÁZISOK ELMÉLETE 7. ELŐADÁS

név	város	utca
Gálvölgyi J.	Budapest	Nyereg u. 2.
Gálvölgyi J.	Budapest	Kantár u. 3.
⋮		

Kulcs: elromlik az ODL-es kulcs, lehet, hogy ami ott kulcs volt, itt már nem lesz az
⇒ baj