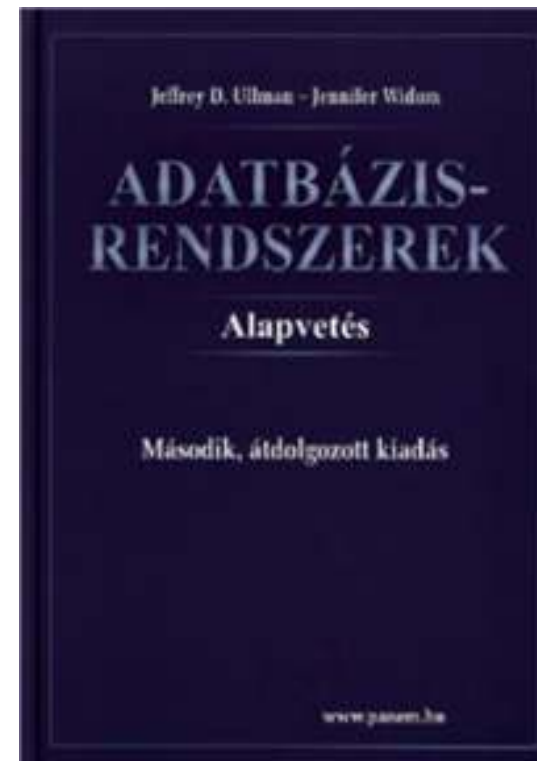


Relációs algebra 5.rész kiterjesztés

Tankönyv: Ullman-Widom:
Adatbázisrendszerek Alapvetés
Második, átdolgozott kiadás,
Panem, 2009

5.2. Kiterjesztett műveletek a relációs algebrában

Megj.: Ez a Relációs algebra 3.rész
02C_RelAlg3multihz.pdf folytatása



Relációs algebra kibővítése

- Az eddig tanult műveleteket: **vetítés** (Π), **kiválasztás** (σ), **halmazműveletek**: **unió** (\cup), **különbség** ($-$), **metszet** (\cap), **szorzás**: **természetes összekapcsolás** (\bowtie), **direkt-szorzat** (\times), stb. **multihalmazok fölött értelmezzük, mint az SQL-ben**, egy reláció nem sorok halmazából, hanem **multihalmazából áll**, vagyis megengedett a sorok ismétlődése.
- Ezekon kívül a **SELECT kiegészítéseinek és záradékainak** megfelelően **új műveletekkel is kibővítjük** a rel. algebrát:
 - **Ismétlődések megszüntetése** (δ) - **select distinct ..**
 - **Vetítési művelet kiterjesztése** (Π_{lista}) - **select kif [as onev]..**
 - **Rendezési művelet** (τ_{lista}) - **order by..**
 - **Összesítő műveletek és csoportosítás** (γ_{lista}) - **group by..**
 - **Külső összekapcsolások** ($\overset{\circ}{\bowtie}$) - **[left | right | full] outer join**

Ismétlődések megszüntetése (duplikátumok kiszűrése)

- **Ismétlődések megszüntetése**: $R1 := \delta(R2)$
- A művelet jelentése: R2 multihalmazból R1 halmazt állít elő, vagyis az R2-ben egyszer vagy többször előforduló sorok csak egyszer szerepelnek az R1-ben.
- A DISTINCT reprezentálására szolgál
- A δ speciális esete lesz az általánosabb γ műveletnek

$R =$ (

A	B
1	2
3	4
1	2

$\delta(R) =$

A	B
1	2
3	4

Összesítő (aggregáló) függvények

R =

A	B
1	3
3	4
3	2

$$\text{SUM}(A) = 7$$

$$\text{COUNT}(A) = 3$$

$$\text{MIN}(B) = 2$$

$$\text{MAX}(B) = 4$$

$$\text{AVG}(B) = 3$$

Csoportosítás és összesítés $\gamma_L(R)$

- A művelet a csoportosítást (**GROUP BY**), a csoportokon végezhető **összesítő függvényeket** (AVG, SUM, COUNT, MIN, MAX) reprezentálja.
- A művelet jele: $\gamma_L(R)$.
- Itt az **L** lista valamennyi eleme a következők egyike:
 - R egy attribútuma: ez az attribútum egyike a csoportosító attribútumoknak (a GROUP BY után jelenik meg).
 - a reláció egyik attribútumára (ez az összesítési attribútumra) alkalmazott összesítő operátor, ha az összesítés eredményére névvel szeretnénk hivatkozni, akkor nyilat és új nevet használunk.

Csoportosítás és összesítés $\gamma_L(R)$

- Osszuk R sorait csoportokba. Egy csoport azokat a sorokat tartalmazza, amelyek az L listán szereplő csoportosítási attribútumokhoz tartozó értékei megegyeznek
- Vagyis ezen attribútumok minden egyes különböző értéke egy csoportot alkot.
- Minden egyes csoporthoz számoljuk ki az L lista összesítési attribútumaira vonatkozó összesítéseket
- Az eredmény minden egyes csoportra egy sor:
 1. A csoportosítási attribútumok és
 2. Az összesítési attribútumra vonatkozó összesítések (az adott csoport összes sorára)

Példa: Csoportosításra és összesítésre

R =

A	B	C
1	2	3
4	5	6
1	2	5

$\gamma_{A,B,AVG(C)} \rightarrow X (R) = ??$

Először csoportosítunk

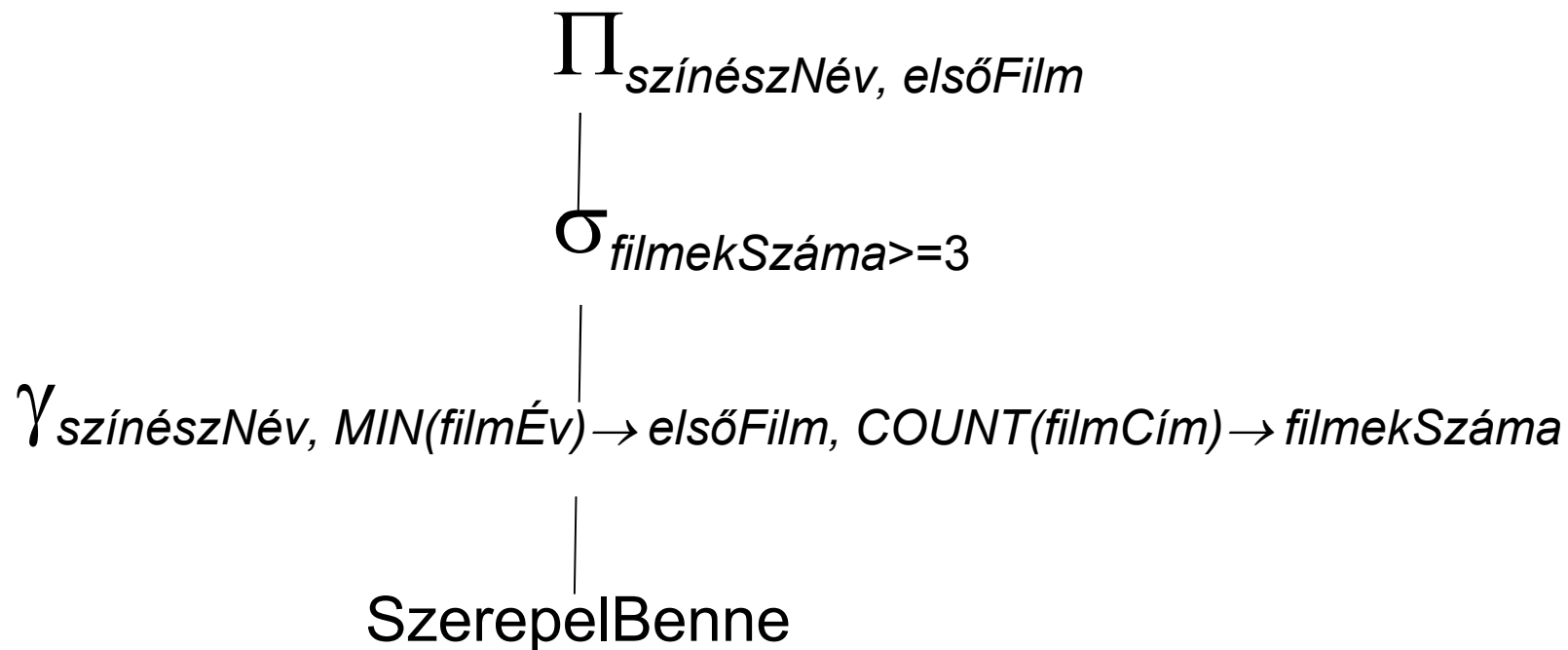
A	B	C
1	2	3
1	2	5
4	5	6

majd csoportonként
összesítünk:

A	B	X
1	2	4
4	5	6

Tankönyv példája

- Adjuk meg azokat a színészeket, akik már szerepeltek legalább három filmben illetve adjuk meg az első olyan filmet is, amiben szerepelt:



Külső összekapcsolások

- Ez **nem relációs algebrai művelet**, uis kilép a modellből.
- Lehet baloldali, jobboldali, teljes külső összekapcsolás.
- R, S sémái $R(A_1, \dots, A_n, B_1, \dots, B_k)$, ill. $S(B_1, \dots, B_k, C_1, \dots, C_m)$
- $R \overset{\circ}{\bowtie} S = R \bowtie S$ relációt kiegészítjük az R és S soraival, a hiányzó helyekre NULL értéket írva megőrzi a „lógó sorokat”
- Van teljes, baloldali és jobboldali külső összekapcsolás attól függően, hogy melyik oldalon szereplő reláció sorait adjuk hozzá az eredményhez (a lógó sorokat kiegészítve NULL értékkel) \perp szimbólummal.

Példák külső összekapcsolásokra

A	B	C
1	2	3
4	5	6
7	8	9

R reláció

B	C	D
2	3	10
2	3	11
6	7	12

S reláció

A	B	C	D
1	2	3	10
1	2	3	11
4	5	6	⊥
7	8	9	⊥
⊥	6	7	12

$R \bowtie^{\circ} S$ eredmény

A	B	C	D
1	2	3	10
1	2	3	11
4	5	6	⊥
7	8	9	⊥

$R \bowtie^{\circ}_L S$ eredmény

A	B	C	D
1	2	3	10
1	2	3	11
⊥	6	7	12

$R \bowtie^{\circ}_R S$ eredmény