

Relációs algebra

Gugolya László

Relációs algebra

Relációs algebra: adattáblákon végzett matematikai műveletek rendszere.

- Az adatbázisok lekérdezésének elméleti alapját képezi, azaz a lekérdező nyelvek implementálják őket (pl. SQL)
- A műveletek operandusai relációk
- A műveletek eredményei relációk (nulla sort is tartalmazhat)
- A relációkat halmaznak tekintjük, melynek elemei a reláció sorai
- Halmaz orientált
- Algebrai eszközökkel dolgozik
- Procedurális, műveletekkel adjuk meg mi történjen (hogyan)

Relációs algebra műveletei

- Szelekció - sorok kiválasztása a relációból
- Projekció - oszlopok kiválasztása a relációból
- Descartes szorzat két reláció kombinációja
- Összekapcsolás (levezethető)
- Különbség
- Unió
- Metszet (levezethető)
- Osztás (levezethető)

Ismétlés

Reláció, vagy adattábla az $R(A_1, \dots, A_n)$ relációs séma felett:

$$T \subseteq \text{dom}(A_1) \times \dots \times \text{dom}(A_n).$$

Példa: a Könyv (könyvszám, szerző, cím) relációséma feletti adattábla:

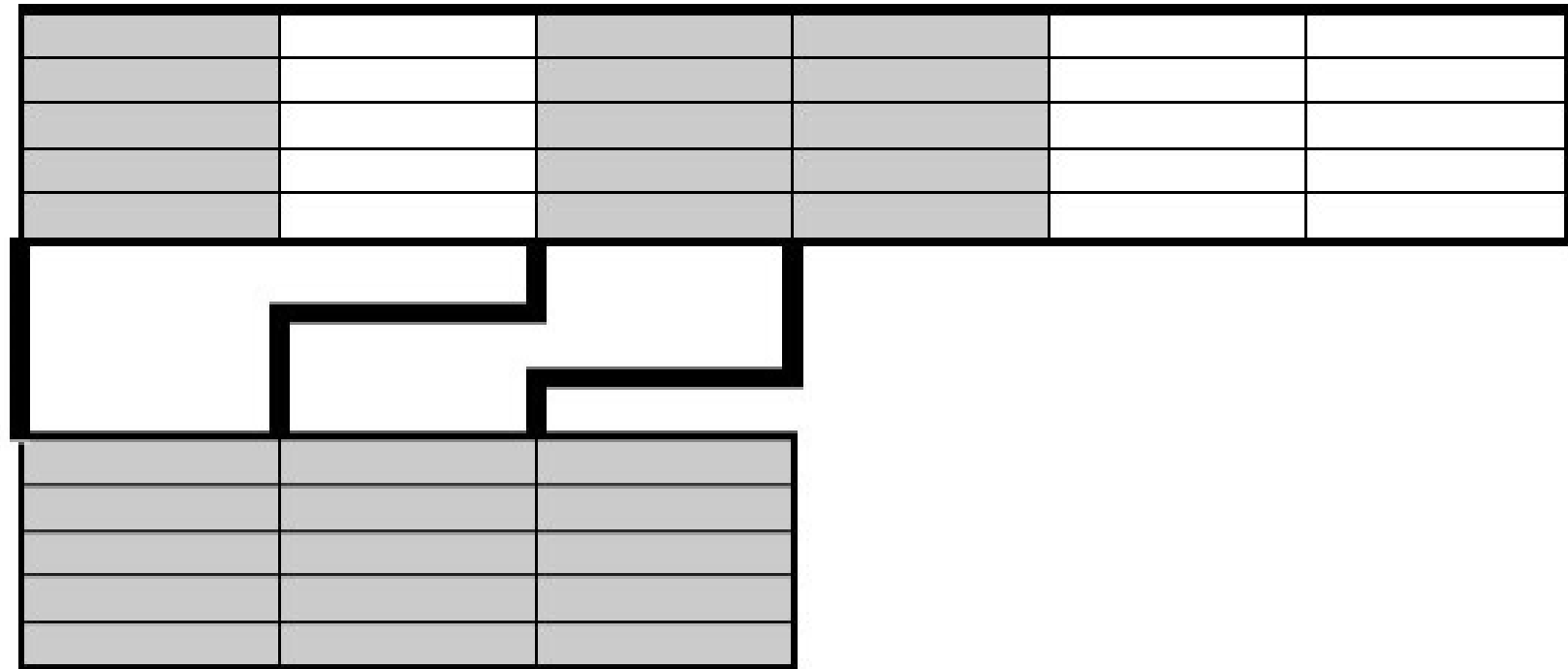
$$T \subseteq \text{dom}(\text{könyvszám}) \times \text{dom}(\text{szerző}) \times \text{dom}(\text{cím}):$$

***Az adattáblát sorok halmazának tekintjük.
A tábla minden sora különböző.***

Redukciós műveletek

Projekció (vetítés): oszlopok kiválasztása, egy operandusú művelet, vertikális megszorítás

Jelölés: $\pi_{\text{attribútumlista}}(\text{tábla})$



Redukciós műveletek

Példa: **Könyv**

$\pi_{\text{szerző,cím}}(\text{Könyv})$

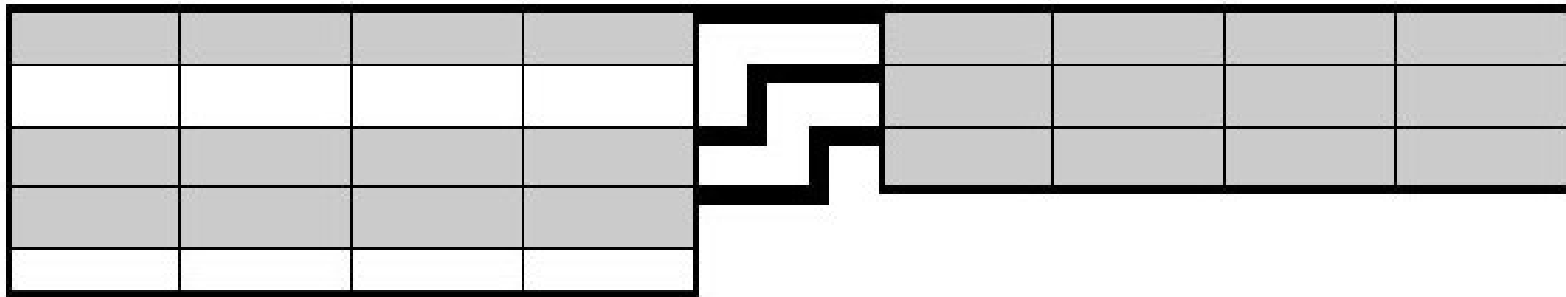
<i>K.szám</i>	<i>Szerző</i>	<i>Cím</i>
1121	Sályi	Adatbázisok
3655	Radó	Világatlasz
2276	Karinthy	Így írtok ti
1782	Jókai	Aranyember
3140	Karinthy	Így írtok ti

<i>Szerző</i>	<i>Cím</i>
Sályi	Adatbázisok
Radó	Világatlasz
Karinthy	Így írtok ti
Jókai	Aranyember

Megjegyzés: a sorok száma csökkenhet, ha az attribútumlista nem tartalmazza a kulcsot. Így az eredményben azonos sorok is előfordulhatnak, ezeket meg kell szüntetni, miért? Implementációja automatikusan nem szünteti meg a duplikált sorokat, de lehetőséget biztosít rá(pl.: SQL - distinct)

Szelekció (kiválasztás): sorok kiválasztása,
horizontális megszorítás, egy operandusú
művelet

Jelölés: $\sigma_{\text{feltétel}}(\text{tábla})$



A szelekció kommutatív:

$$\sigma_{f_1}(\sigma_{f_2}(\text{tábla})) = \sigma_{f_2}(\sigma_{f_1}(\text{tábla})) = \sigma_{(f_1 \text{ AND } f_2)}(\text{tábla})$$

Szelekció

A megszorítást egy logikai feltétellel adhatjuk meg

A feltételben

- a reláció attribútumai
- a reláció műveletek (<, >, =, !=, >=, <=)
- logikai műveletek (not \neg , and \wedge , or \vee)

Példa: Könyv

<i>K.szám</i>	<i>Szerző</i>	<i>Cím</i>	<i>O.szám</i>	<i>Kivétel</i>
1121	Sályi	Adatbázisok	355	2006.03.15
3655	Radó	Világatlasz	122	2007.07.12
2276	Karinthy	Így írtok ti		
1782	Jókai	Aranyember	355	2006.09.23

$\sigma_{\text{kivétel} < 2007.01.01}$ **(Könyv)**

<i>K.szám</i>	<i>Szerző</i>	<i>Cím</i>	<i>O.szám</i>	<i>Kivétel</i>
1121	Sályi	Adatbázisok	355	2006.03.15
1782	Jókai	Aranyember	355	2006.09.23

Halmazműveletek

Az $R_1(A_1, \dots, A_n)$ és $R_2(B_1, \dots, B_m)$ relációsémák **kompatibilisek**, ha $n = m$ és $\text{dom}(A_i) = \text{dom}(B_i)$ minden i -re.

Példa: két könyvtár állománya:

Könyv1 (szerzőnév, cím), Könyv2 (szerző,
könyvcím)

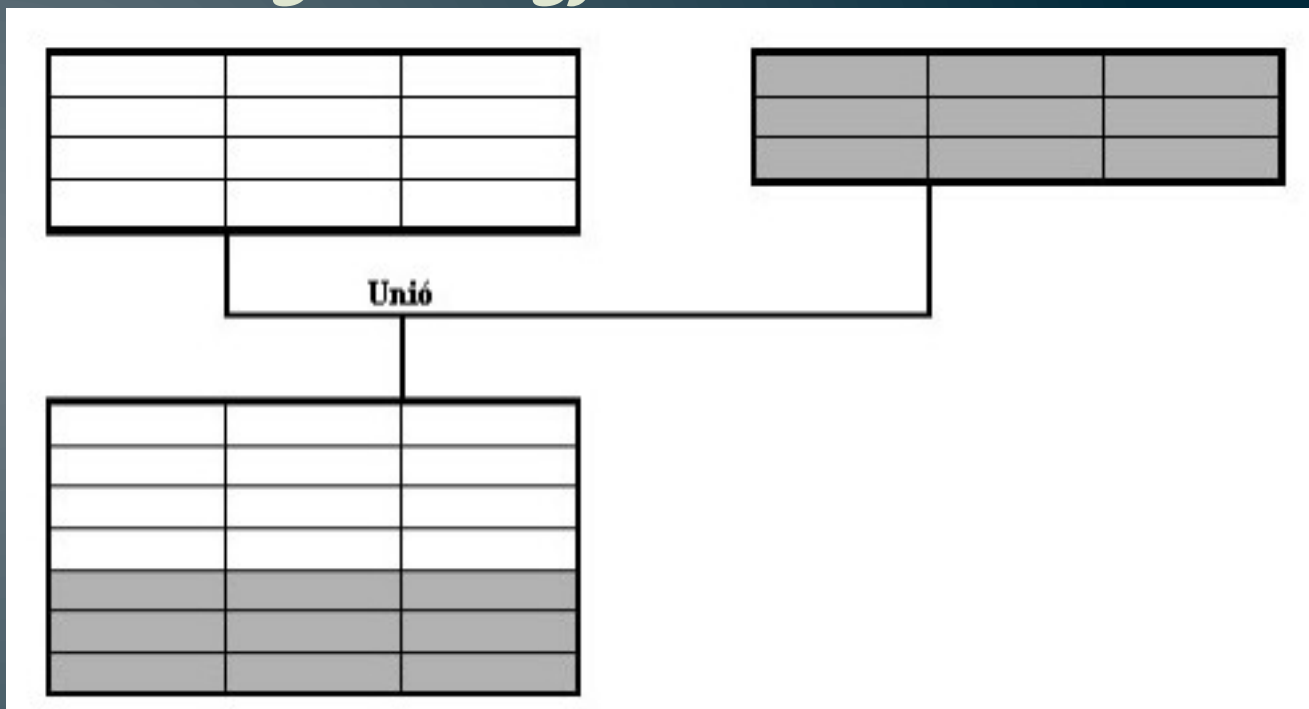
$\text{dom}(\text{szerzőnév}) = \text{dom}(\text{szerző})$

$\text{dom}(\text{cím}) = \text{dom}(\text{könyvcím})$

A halmazműveleteket kompatibilis sémák feletti táblákon értelmezzük.

Unió

A két reláció legalább egyikében előforduló sorok



A művelet végrehajtása:

- a két tábla egymás után írása,
- ismétlődő sorok kiszűrése.

Unió - Példa

Könyv1 tábla:

Sályi	Adatbázisok
Jókai	Aranyember
Radó	Világatlasz

Könyv2 tábla:

Jókai	Aranyember
Karinthy	Így írtok ti

Könyv1 U Könyv2 tábla:

Sályi	Adatbázisok
Jókai	Aranyember
Radó	Világatlasz
Karinthy	Így írtok ti

Megjegyzések:

- Az unió kommutatív művelet.
- Csak kompatibilis táblák unióját lehet képezni.

Példa: Dolgozó (adószám, név, beosztás, fizetés)
Ügyfél (adószám, név, lakcím)

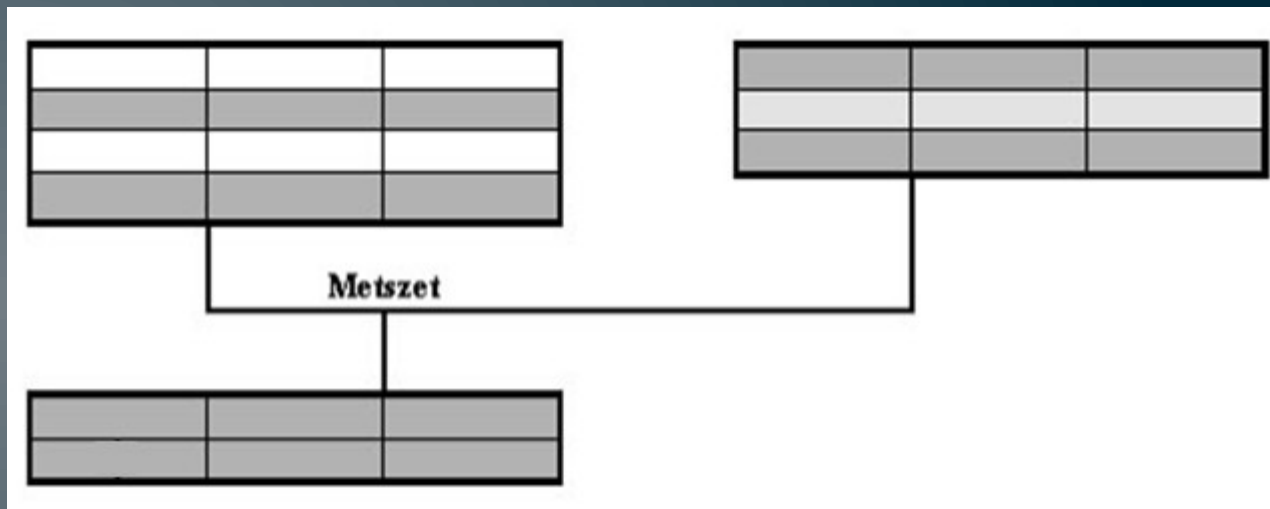
Hibás: Személy = Dolgozó U Ügyfél

Helyes:

$$\text{Személy} = \pi_{\text{adószám, név}}(\text{Dolgozó}) \cup \pi_{\text{adószám, név}}(\text{Ügyfél})$$

Metszet

A két reláció mindegyikében előforduló sorok



Megjegyzések:

- A metszet kommutatív művelet.
- Csak kompatibilis táblák metszetét lehet képezni.

Példa: $\pi_{\text{adószám, név}}(\text{Dolgozó}) \cap \pi_{\text{adószám, név}}(\text{Ügyfél})$

Metszet - Példa

Könyv1 *tábla:*

Sályi	Adatbázisok
Jókai	Aranyember
Radó	Világatlasz

Könyv2 *tábla:*

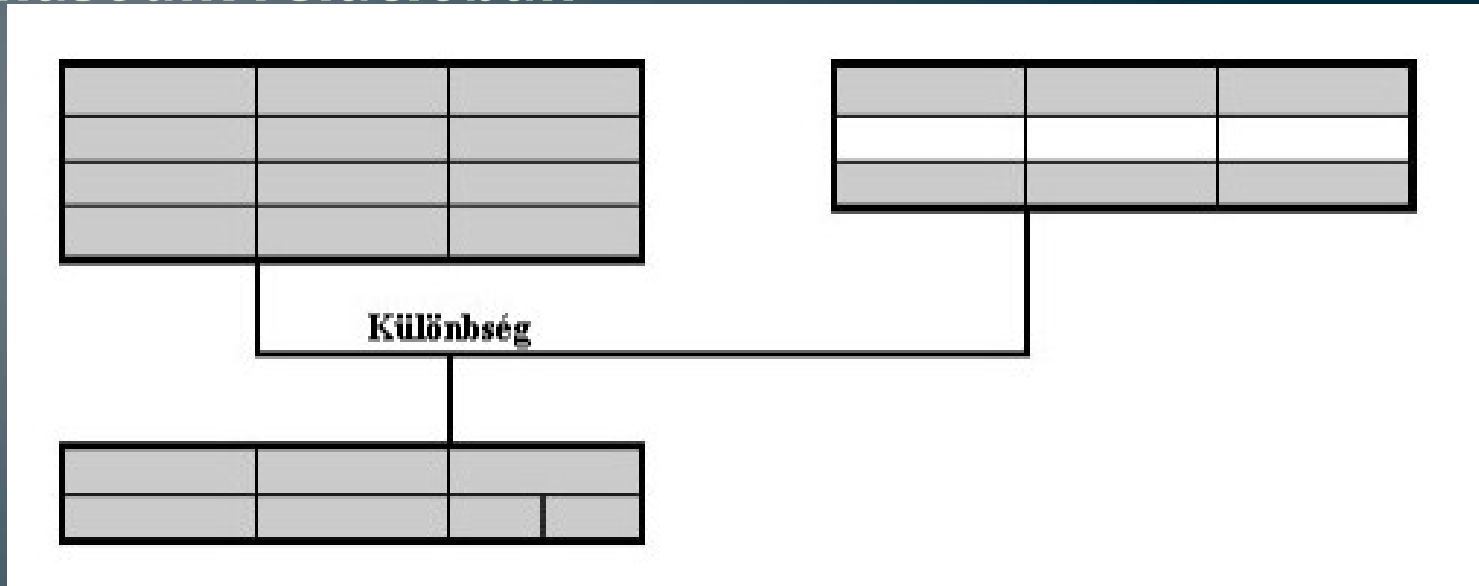
Jókai	Aranyember
Karinthy	Így írtok ti

Könyv1 \cap Könyv2 *tábla:*

Jókai	Aranyember
-------	------------

Különbség

Az első reláció azon sorai, melyek nem szerepelnek a második relációban



Megjegyzések:

- A különbség nem kommutatív művelet.
- Csak kompatibilis táblák különbségét lehet képezni.

Különbség - Példa

Könyv1 tábla:

Sályi	Adatbázisok
Jókai	Aranyember
Radó	Világatlasz

Könyv2 tábla:

Jókai	Aranyember
Karinthy	Így írtok ti

Könyv1 – Könyv2 tábla:

Sályi	Adatbázisok
Radó	Világatlasz

Kombinációs műveletek

Descartes-szorzat

Adott: $R_1(A_1, \dots, A_n), R_2(B_1, \dots, B_m)$

$$T_1 \subseteq \text{dom}(A_1) \times \dots \times \text{dom}(A_n),$$

$$T_2 \subseteq \text{dom}(B_1) \times \dots \times \text{dom}(B_m)$$

Descartes-szorzat: $R(A_1, \dots, A_n, B_1, \dots, B_m)$ séma felett

$$T \subseteq \text{dom}(A_1) \times \dots \times \text{dom}(A_n) \times \text{dom}(B_1) \times \dots \times$$

$\text{dom}(B_m)$ tábla: ***T_1 minden sorát párosítjuk T_2 minden sorával.***

Jele: $T = T_1 \times T_2$

Descartes-szorzat

- Két reláció rekordjainak minden kombinációban történő összepárosítása
- Az eredményben azonos nevű attribútumok is lehetnek, ezért átnevezésre is szükség lehet

1		
2		
3		
4		

1			
2			
3			

1			1			
2			2			
3			3			
4			1			
1			2			
2			3			
3			1			
4			2			
1			3			
2			1			
3			2			
4			3			

Példa Descartes-szorzatra:

T_1 : A_1 A_2 A_3

a b c

b d e

f c b

T_2 : B_1 B_2

x y

y z

$T_1 \times T_2$: A_1 A_2 A_3 B_1 B_2

a b c x y

a b c y z

b d e x y

b d e y z

f c b x y

f c b y z

Tulajdonságok:

- Ha T_1 és T_2 sorainak száma r_1 ill. r_2 , oszlopainak száma c_1 és c_2 ,
akkor a T táblában $r_1 \cdot r_2$ sor és $c_1 + c_2$ oszlop van.
- Ha $T = T_1 \times T_2$, akkor projekcióval visszakaphatók az eredeti

Természetes összekapcsolás (Natural join)

Példa:

Könyv (könyvszám, szerző, cím, *olvasószám*, kivétel)

Olvasó (olvasószám, név, lakcím)

A kikölcsönzött könyvek listája az olvasók adataival:

Kolv (könyvszám, szerző, cím, olvasószám, kivétel, név, lakcím)

Kolv = Könyv * Olvasó

A természetes összekapcsolás definíciója

A és B attribútumhalmazok, $R_1(A)$ és $R_2(B)$ sémák,

T_1 és T_2 táblák a sémák felett.

$X = A \cap B$ nem üres.

T_1 és T_2 természetes összekapcsolása egy $R(A \cup B)$ feletti T tábla:

$$T = \pi_{A \cup B}(\sigma_{R1.X=R2.X}(T_1 \times T_2))$$

Jelölés: $T = T_1 * T_2$

Összekapcsolás: *kiindulási táblák*

A Könyv tábla:

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>
1121	Sályi	Adatbázisok		
3655	Radó	Világatlasz	122	2004.07.12
2276	Karinthy	Így írtok ti		
1782	Jókai	Aranyember	355	2004.09.23

Az Olvasó tábla:

<i>Oszám</i>	<i>Név</i>	<i>Lakcím</i>
122	Kiss István	Szeged, Virág u. 10.
612	Nagy Ágnes	Szentes, Petőfi út 38.
355	Tóth András	Budapest, Jég u. 3.

Összekapcsolás 1. lépés: *Descartes szorzat*

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Oszám</i>	<i>Név</i>	<i>Lakcím</i>
1121	Sályi	Adatbázisok			122	Kiss	Szeged...
3655	Radó	Világatlasz	122	2004.07.12	122	Kiss	Szeged...
2276	Kar.	Így írtok ti			122	Kiss	Szeged...
1782	Jókai	Aranyember	355	2004.09.23	122	Kiss	Szeged...
1121	Sályi	Adatbázisok			612	Nagy	Szentes...
3655	Radó	Világatlasz	122	2004.07.12	612	Nagy	Szentes...
2276	Kar.	Így írtok ti			612	Nagy	Szentes...
1782	Jókai	Aranyember	355	2004.09.23	612	Nagy	Szentes...
1121	Sályi	Adatbázisok			355	Tóth	
		Budapest...					
3655	Radó	Világatlasz	122	2004.07.12	355	Tóth	Budapest...
2276	Kar.	Így írtok ti			355	Tóth	Budapest...
1782	Jókai	Aranyember	355	2004.09.23	355	Tóth	Budapest...

Összekapcsolás 2. lépés: *Szelekció*

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Oszám</i>	<i>Név</i>
		<i>Lakcím</i>				
3655	Radó	Világatlasz Szeged...	122	2004.07.12	122	Kiss
1782	Jókai	Aranyember Budapest...	355	2004.09.23	355	Tóth

Összekapcsolás 3. lépés: *Projekció*

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Név</i>	<i>Lakcím</i>
3655	Radó	Világatlasz	122	2004.07.12	Kiss	Szeged...
1782	Jókai	Aranyember	355	2004.09.23	Tóth	Budapest...

A Könyv tábla:

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>
1121	Sályi	Adatbázisok		
3655	Radó	Világatlasz	122	2004.07.12
2276	Karinthy	Így írtok ti		
1782	Jókai	Aranyember	355	2004.09.23

Az Olvasó tábla:

<i>Oszám</i>	<i>Név</i>	<i>Lakcím</i>
122	Kiss István	Szeged, Virág u. 10.
612	Nagy Ágnes	Szentes, Petőfi út 38.
355	Tóth András	Budapest, Jég u. 3.

A Kolv = Könyv*Olvasó tábla:

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Név</i>	<i>Lakcím</i>
3655	Radó	Világatlasz	122	2004.07.12	Kiss	Szeged, Virág u.10
1782	Jókai	Aranyember	355	2004.09.23	Tóth	Budapest, Jég u.3.

Megjegyzés a természetes összekapcsoláshoz:

Ha az $R_1(A)$ és $R_2(B)$ sémák feletti T_1 és T_2 táblákra $T = T_1 * T_2$, akkor projekcióval általában nem kaphatók vissza az eredeti táblák:

$$\pi_A(T) \neq T_1$$

$$\pi_B(T) \neq T_2.$$

Ennek oka: lógó sorok (dangling rows), amelyek nem találnak párt maguknak.

Külső összekapcsolás (Outer join)

Bal oldali külső összekapcsolás: $T_1 (+)^ T_2$*

*Jobb oldali külső összekapcsolás: $T_1 *(+) T_2$*

Kétoldali külső összekapcsolás: $T_1 (+)^(+) T_2$*

A Könyv tábla:

<i>Kszám</i>	<i>szerző</i>	<i>cím</i>	<i>olvasószám</i>	<i>kivétel</i>
1121	Sályi	Adatbázisok		
3655	Radó	Világatlasz	122	2004.07.12
2276	Karinthy	Így írtok ti		
1782	Jókai	Aranyember	355	2004.09.23

Az Olvasó tábla:

<i>Oszám</i>	<i>név</i>	<i>lakcím</i>
122	Kiss István	Szeged, Virág u. 10.
612	Nagy Ágnes	Szentes, Petőfi út 38.
355	Tóth András	Budapest, Jég u. 3.

A Könyv (+)* Olvasó tábla:

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Név</i>	<i>Lakcím</i>
1121	Sályi	Adatbázisok				
3655	Radó	Világatlasz	122	2004.07.12	Kiss	Szeged, Virág u.10
2276	Karinthy	Így írtok ti				
1782	Jókai	Aranyember	355	2004.09.23	Tóth	Budapest, Jég u.3.

A Könyv tábla:

<i>Kszám</i>	<i>szerző</i>	<i>cím</i>	<i>olvasószám</i>	<i>kivétel</i>
1121	Sályi	Adatbázisok		
3655	Radó	Világatlasz	122	2004.07.12
2276	Karinthy	Így írtok ti		
1782	Jókai	Aranyember	355	2004.09.23

Az Olvasó tábla:

<i>Oszám</i>	<i>név</i>	<i>lakcím</i>
122	Kiss István	Szeged, Virág u. 10.
612	Nagy Ágnes	Szentes, Petőfi út 38.
355	Tóth András	Budapest, Jég u. 3.

A Könyv *(+) Olvasó tábla:

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Név</i>	<i>Lakcím</i>
3655	Radó	Világatlasz	122	2004.07.12	Kiss	Szeged, Virág u.10
			612		Nagy	Szentes, Petőfi út 38
1782	Jókai	Aranyember	355	2004.09.23	Tóth	Budapest, Jég u.3.

A Könyv tábla:

<i>Kszám</i>	<i>szerző</i>	<i>cím</i>	<i>olvasószám</i>	<i>kivétel</i>
1121	Sályi	Adatbázisok		
3655	Radó	Világatlasz	122	2004.07.12
2276	Karinthy	Így írtok ti		
1782	Jókai	Aranyember	355	2004.09.23

Az Olvasó tábla:

<i>Oszám</i>	<i>név</i>	<i>lakcím</i>
122	Kiss István	Szeged, Virág u. 10.
612	Nagy Ágnes	Szentes, Petőfi út 38.
355	Tóth András	Budapest, Jég u. 3.

A Könyv (+)*(+) Olvasó tábla:

<i>Kszám</i>	<i>Szerző</i>	<i>Cím</i>	<i>Oszám</i>	<i>Kivétel</i>	<i>Név</i>	<i>Lakcím</i>
1121	Sályi	Adatbázisok				
3655	Radó	Világatlasz	122	2004.07.12	Kiss	Szeged, Virág u.10
2276	Karinthy	Így írtok ti				
			612		Nagy	Szentes, Petőfi út 38
1782	Jókai	Aranyember	355	2004.09.23	Tóth	Budapest, Jég u.3.

Megjegyzés a külső összekapcsoláshoz

Projekcióval visszakaphatók az eredeti táblák:

Ha az $R_1(A)$ és $R_2(B)$ sémák feletti T_1 és T_2 táblákra

$$T = T_1 (+)^* T_2, \text{ akkor } \pi_A(T) = T_1,$$

$$T = T_1 *(+) T_2, \text{ akkor } \pi_B(T) = T_2,$$

$$T = T_1 (+)^*(+) T_2, \text{ akkor } \pi_A(T) = T_1 \text{ és } \pi_B(T) = T_2.$$

Egy gyakorlati példa

Egyesítendő nyilvántartások:

K1 (ISBN, szerző, cím, kiadó)

K2 (ISBN, szerző, cím, év, ár)

1. megoldás: $K = \pi_{\text{ISBN, szerző, cím}}(K1) \cup \pi_{\text{ISBN, szerző, cím}}(K2)$

K (ISBN, szerző, cím)

2. megoldás: $K = K1 (+)^*(+) (K2)$

K (ISBN, szerző, cím, kiadó, év, ár)

3. megoldás: $K = K1 (+)^*(+) \pi_{\text{ISBN, év, ár}}(K2)$

K (ISBN, szerző, cím, kiadó, év, ár)

Théta-összekapcsolás (theta-join)

$$T = \sigma_{\text{feltétel}}(T_1 \times T_2)$$

$$\text{Jelölés: } T = T_1 *_{\text{feltétel}} T_2$$

Példa:

Raktár (raktárkód, mennyiség)


Vevő (vevőkód, igény)

Ajánlat = Raktár $*_{\text{igény} \leq \text{mennyiség}}$ Vevő

Ajánlat (raktárkód, mennyiség, vevőkód, igény)

Műveleti sorrend

Mivel több művelet is elvégezhető, így ezeknek definiálni kell az elvégzési sorrendjét, azaz a műveleti precedenciát:

- 1. szelekció, projekció, [σ , π]
- 2. Descartes-szorzat, kapcsolat(join) ,
- 3. metszet, \cap
- 4. unió, különbség, [\cup , $-$]

Multihalmazok

Halmaz: {4, 1, 3}

Pl. adattábla minden sora különböző.

Multihalmaz: {4, 1, 4, 3, 1, 2}

Pl. adattáblában azonos sorok lehetnek.

A relációs algebra multihalmazokra is kidolgozható.

DBMS-ek multihalmazokkal dolgoznak:

- Fizikai adattárolás
- Sebesség
- Gyakorlati igény is szükség lehet, például
Dolgozó (név, adószám, lakcím, fizetés)
 $\text{Dolg} = \pi_{\text{név}, \text{fizetés}}(\text{Dolgozó})$