Adatbázisok 1. Relációs adatbázis tervezés – 3. rész

Funkcionális függőségek

Felbontások

Normálformák

Az összes következmény FF megtalálása

- Motiváció: „normalizálás”, melynek során egy relációsémát több sémára bonthatunk szét.
- Példa: *ABCD* FF-k: *AB* ->*C*, *C* ->*D* és

  D ->*A*.
  - Bontsuk fel *ABC* és *AD*-re. Milyen FF-k teljesülnek *ABC* –n?
  - Nem csak *AB* ->*C*, de *C* ->*A* is!

Miért?

ABCD $a_1 b_1 c d_1$ $a_2 b_2 c d_2$

ebből $\downarrow$ $\downarrow$

ABC $a_1 b_1 c$ $a_2 b_2 c$

$d_1 = d_2$ mert
C -> D

$a_1 = a_2$ mert
D -> A

Emiatt, ha két vetített sor C-n megegyezik, akkor A-n is, azaz: *C -> A*.

Alapötlet

1. Induljunk ki a megadott FF-ekből és keressük meg az összes *nem triviális* FF-t, ami a megadott FF-ekből következik.
    - Nem triviális = a jobboldalt nem tartalmazza a bal.
2. Csak azokkal az FF-kel foglalkozzunk, amelyekben a projektált séma attribútumai szerepelnek.

Exponenciális algoritmus

1. Minden X attribútumhalmazra számítsuk ki X^+-t.
2. Adjuk hozzá a függőségeinkhez $X \to A$-t minden A-ra $X^+ - X$-ből.
3. Dobjuk ki $XY \to A$-t, ha $X \to A$ is teljesül.
   - Mert $XY \to A$ az $X \to A$-ból minden esetben következik.
4. Végül csak azokat az FF-ket használjuk, amelyekben csak a projektált attribútumok szerepelnek.

Néhány trükk

- Az üreshalmaznak és az összes attribútum halmazának nem kell kiszámolni a lezártját.
- Ha X^+ = az összes attribútum, akkor egyetlen X- t tartalmazó halmaznak sem kell kiszámítani a lezártját.

Példa: FF-k projekciója

- ABC, $A \to B$ és $B \to C$ FF-kel. Projektáljunk AC-re.
  - $A^+ = ABC$; ebből $A \to B$, $A \to C$.
    - Nem kell kiszámítani AB^+ és AC^+ lezárásokat.
  - $B^+ = BC$; ebből $B \to C$.
  - $C^+ = C$; semmit nem ad.
  - $BC^+ = BC$; semmit nem ad.

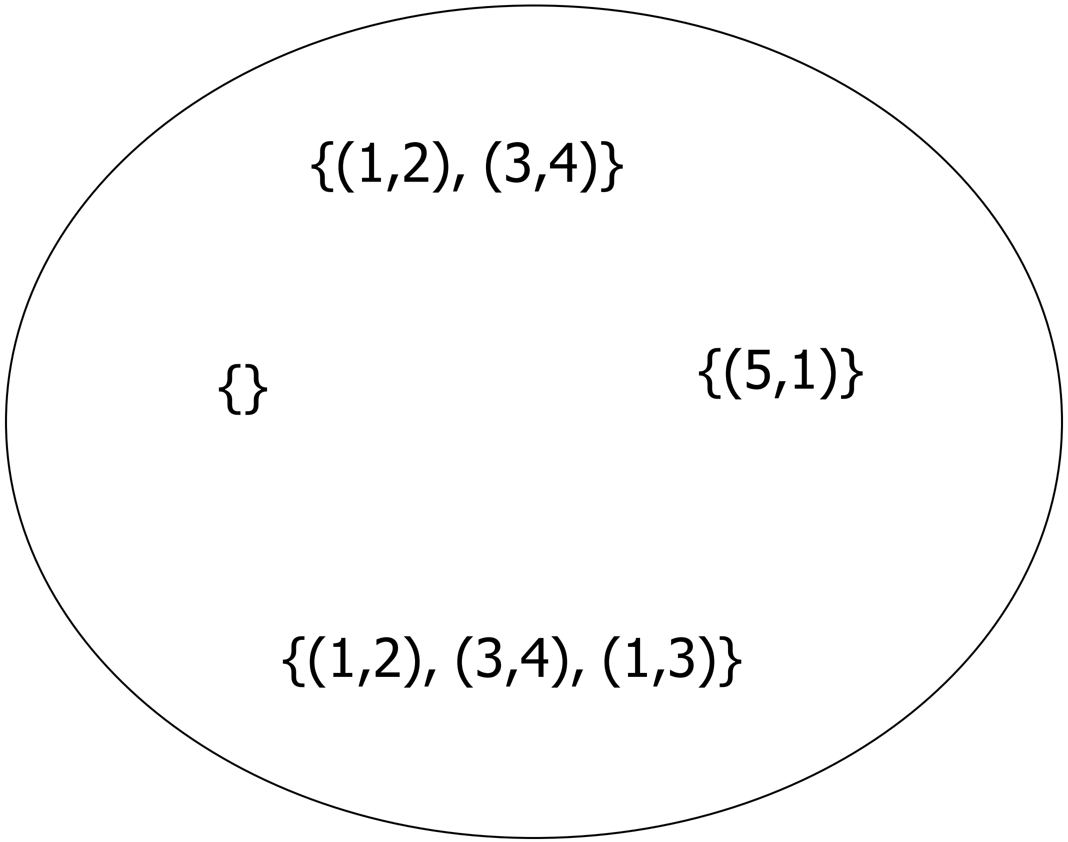
Példa -- folytatása

- A kapott FF-ek: $A \to B$, $A \to C$ és $B \to C$.
- AC -re projekció: $A \to C$.

Az FF-k geometriai reprezentációja

- Vegyük egy reláció összes lehetséges *előfordulásainak* halmazát.
- Azaz az összes olyan sorhalmazt, mely sorok komponensei „megfelelőek”.
- Minden ilyen halmaz egy pont a térben.

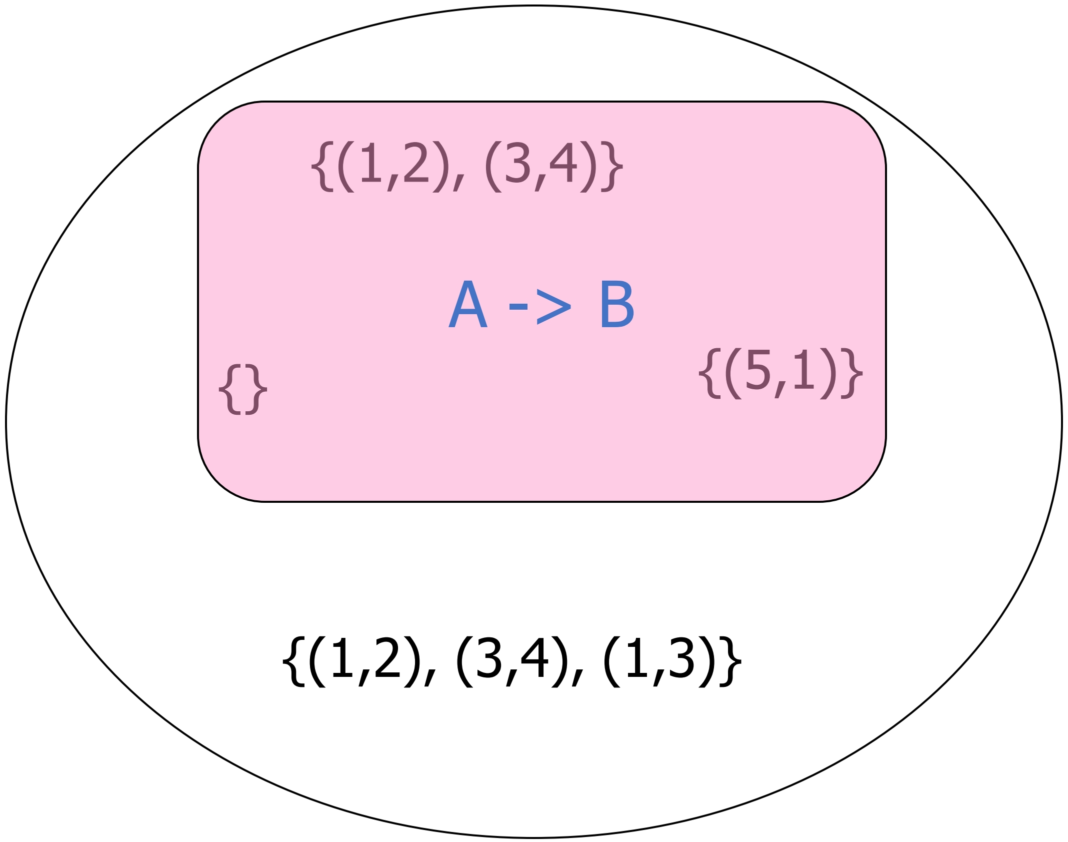
Példa: R(A,B)



Egy FF az előfordulásoknak egy részhalmaza

- Minden *X -> A* FF megadható azon előfordulások részhalmazaként, mely teljesíti FF-t.
- Így minden FF egy régióval jellemezhető a térben
- A triviális FF-k azok, melyeknél ez a régió a teljes tér.
  - Példa: *A -> A*.

Példa: A -> B R(A,B) fölött

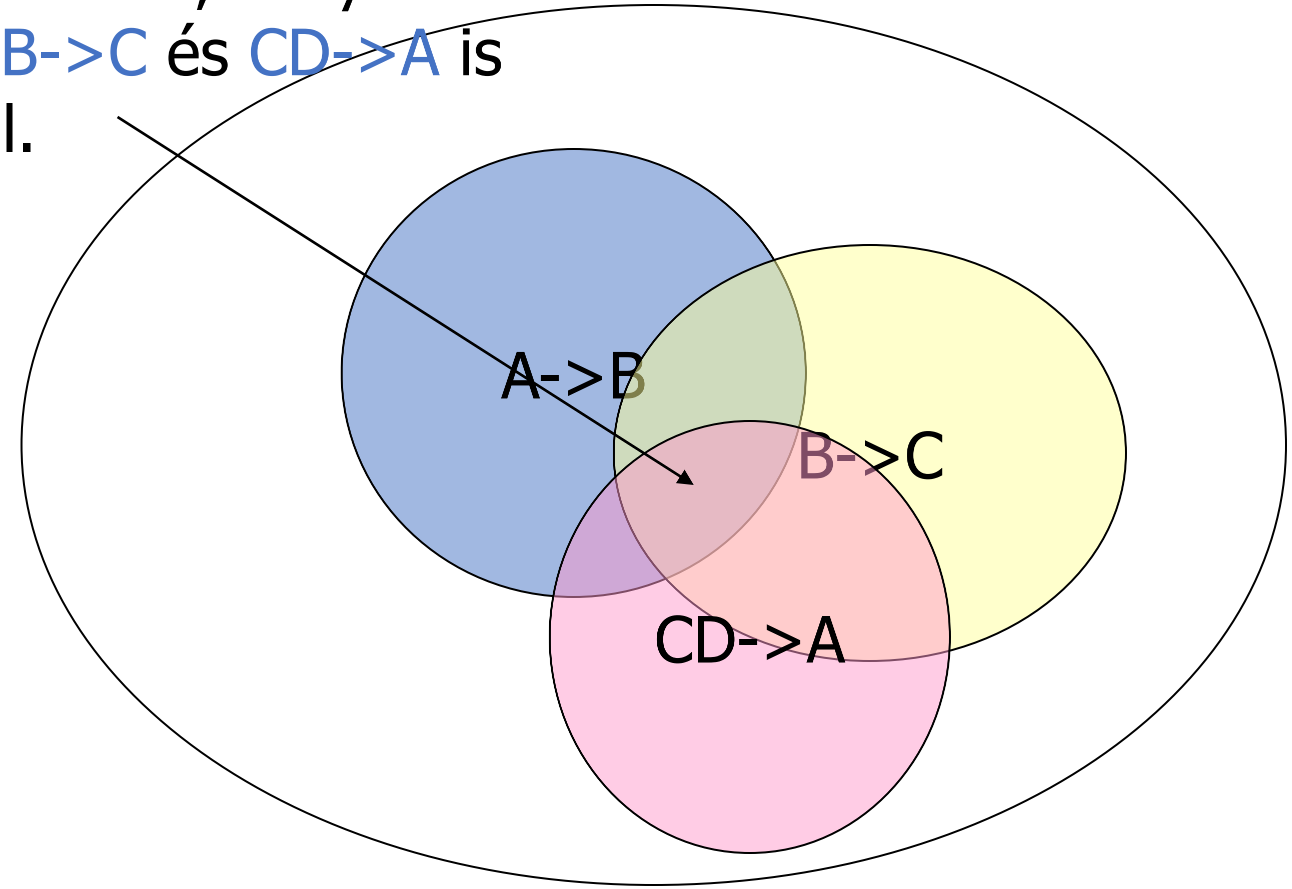


FF-k halmazának reprezentálása

- Ha egy-egy FF előfordulásoknak egy halmazával reprezentálható, akkor az FF-ek halmaza az előbbi halmazok metszetével lesz egyenlő.
  - Azaz a metszet = azon előfordulások, amelyekre mindegyik FF teljesül.

Példa

Előfordulások, melyekre A->B, B->C és CD->A is teljesül.



FF-k következtetése

- Ha $Y \text{ -> } B$ FF következik $X_1 \text{ -> } A_1, \ldots, X_n \text{ -> } A_n$ FF-ekből, akkor az $Y \text{ -> } B$ régiójának tartalmaznia kell az $X_i \text{ -> } A_i$ FF-ekhez tartozó régiók metszetét.
  - Azaz: minden előfordulás, ami teljesíti $X_i \text{ -> } A_i$ -t, $Y \text{ -> } B$ -t is teljesíti.
  - Ugyanakkor ha egy előfordulásra teljesül $Y \text{ -> } B$, $X_i \text{ -> } A_i$ nem feltétlen teljesül.

Példa