# Relációs adatbázisok tervezése

- **E/K diagram transzformálásával kapott relációs adatbázisséma optimalizálása függőségek ismeretében.**
- A **lehetséges előfordulások** azok, amelyek az előre adott **F** megszorításokat, függőségeket kielégítik.
- A sémákhoz hozzávesszük a függőségeket is: **{R(A1,...,An), F}**
- Rossz adatbázisséma:
- **megrendelés(név,cím,áru)** és azok a lehetséges előfordulások, ahol 1 szállítónévhez legfeljebb 1 cím tartozik (**név→cím**).
- Problémák:
  - redundancia: (n,c) párt többször tároljuk.
  - módosítási anomália: ha n címe c'-re módosul, akkor több sorban is változtatni kell.
  - törlési anomália: ha n nem szállít árut, és nullérték nem szerepelhet, akkor az (n,c) információt elveszítjük.
  - beszúrási anomália: hiába ismerjük az n szállító c címét, addig nem tudjuk ezt tárolni, amíg nincs hozzá megrendelés.

| név | cím | áru |
|---|---|---|
| n | c | a1 |
| n | c | a2 |
| n | c | a3 |

# Relációs adatbázisok tervezése

- Megoldás: dekomponálás normálformájú relációkra.
- A **{megrendelés(név,cím,áru), {név→cím}}** sémát dekomponáljuk:
  - **{szállító(név,cím),{név→cím}}**
  - **{szállít(név,áru), ∅}** /* nem adunk meg függőséget */
- A dekomponált sémák már kevésbé redundánsak, és az anomáliákat is kiküszöböltük.

szállító:

| név | cím |
|---|---|
| n1 | c1 |
| n2 | c2 |
| n3 | c3 |

szállít:

| név | áru |
|---|---|
| n1 | a1 |
| n1 | a2 |
| n1 | a3 |
| n2 | a2 |
| n2 | a4 |
| n3 | a1 |
| n3 | a5 |

# Függőségek

- Megszorítások, függőségek típusai:
  - funkcionális függőség (ff)
  - többértékű függőség (tf)
  - gyenge többértékű függőség (gtf)
  - beágyazott többértékű függőség (etf)
  - összekapcsolási függőség (jf)
  - egyenlőséggeneráló függőség (egd)
  - sorgeneráló függőség (tgd)

A definíciók pontos kimondása később fog szerepelni!

- A leggyakoribbak a funkcionális és többértékű függőségek használata.
- A függőségekhez tartozó redundancát csökkentő normálformák:
  - Boyce-Codd normálforma (BCNF)
  - 3. normálforma (3NF)
  - 4. normálforma (4NF)
- Léteznek más normálformák (2NF, 5NF, 6NF), de ezeket gyakorlatban ritkán használjuk. 1NF – en azt értjük, hogy az adatok relációkban szerepelnek, és minden értéket atominak tekintünk, azaz még ha van is belső struktúrája az értéknek, akkor sem vesszük ezt figyelembe, hanem egységként kezeljük.
- A dekompozíciók tulajdonságai:
  - veszteségmentes (a dekomponált táblákban ugyanaz az információ szerepel, mint az eredeti táblában).
  - függőségőrzés (a dekomponált táblákhoz tartozó függőségekből következnek az eredeti táblára felírt függőségek).

# Funkcionális függőségek

- **Definíció**: $X \longrightarrow Y$ funkcionális függőség, ahol $X,Y \subseteq \{A1,A2,\ldots,An\}$.
- **Definíció:** Egy $\{A1,A2,\ldots,An\}$ sémájú **r reláció kielégíti** $X \longrightarrow Y$-t akkor és csak akkor, ha

  $\forall\ t1,t2 \in r$ esetén $(t1[X]=t2[X] \Rightarrow t1[Y]=t2[Y])$.
- A kulcsok megadása is funkcionális függőségek megadását jelenti.
- Ha minden lehetséges relációban $X \rightarrow R$ teljesül, akkor X szuperkulcsa R-nek F-re nézve.
- A minimális szuperkulcsot kulcsnak nevezzük.

# Példa funkcionális függőségekre

- r:

| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 1 |
| 0 | 1 | 1 |

r kielégíti például a következő függőségeket:

$B \rightarrow A$

$B \rightarrow B$

$BC \rightarrow A$

$BC \rightarrow AB$

$BC \rightarrow ABC$

r nem elégíti ki például a következő függőségeket:

$A \rightarrow B$

$B \rightarrow C$

$AB \rightarrow C$

$AB \rightarrow ABC$

# Funkcionális függőségek

- Mennyi funkcionális függőség írható fel n oszlopos séma esetén?
- $2^{2n}$ (Üres halmaz is halmaz!)
- $X \rightarrow \emptyset$ minden relációra teljesül.
- $\emptyset \rightarrow Y$ akkor teljesül, ha a reláció minden sora megegyezik Y-on.
- $X \rightarrow X$ mindig teljesül.
- $Y \subseteq X$ esetén $X \rightarrow Y$ minden relációra teljesül. (Ezek a triviális funkcionális függőségek.)
- Ha megadunk egy F halmazt, amely véges sok funkcionális függőséget tartalmaz, akkor mely függőségeket szerepeltetjük fölöslegesen, olyan értelemben, hogy a többiből már következik a fölösleges függőség?

# Implikációs probléma

- SAT(X $\rightarrow$Y) := { r | r kielégíti X$\rightarrow$Y-t}.
- F={funkcionális függőségek adott halmaza},

  SAT(F):= { r | r kielégíti F összes függőségét}.
- SAT(F)= $\bigcap_{X \rightarrow Y \in F} SAT(X \rightarrow Y)$
- **Logikai implikáció** (adott függőségekből logikailag következik-e egy függőség):

  F|= X$\rightarrow$Y $\Leftrightarrow$ SAT(F) $\subseteq$ SAT(X $\rightarrow$Y).
- Definíció alapján végtelen sok r táblára kellene ellenőrizni.
- Megoldás: Axiomatizálás

# Armstrong-axiómák

Legyen $X,Y \subseteq R$, és XY jelentse az X és Y attribútumhalmazok egyesítését.

- A1 (reflexivitás): $Y \subseteq X$ esetén $X \rightarrow Y$.
- A2 (tranzitivitás): $X \rightarrow Y$ és $Y \rightarrow Z$ esetén $X \rightarrow Z$.
- A3 (bővíthetőség): $X \rightarrow Y$ és tetszőleges Z esetén $XZ \rightarrow YZ$.
- $X \rightarrow Y$ levezethető F-ből, ha **van olyan $X1 \rightarrow Y1$, ..., $Xk \rightarrow Yk$,..., $X \rightarrow Y$ véges levezetés**, hogy $\forall$k-ra
  - $Xk \rightarrow Yk \in F$ vagy
  - $Xk \rightarrow Yk$ az A1,A2,A3 axiómák alapján kapható a levezetésben előtte szereplő függőségekből.

# Armstrong-axiómák

- $F \vdash X \rightarrow Y$, ha $X \rightarrow Y$ levezethető F-ből.
- Mi a kapcsolat a levezethetőség és a logikai implikáció között?
- Az axiómarendszer **helyes,** ha $\forall$ F és $\forall$ $X \rightarrow Y$ esetén

  $F \vdash X \rightarrow Y \;\Rightarrow\; F \models X \rightarrow Y$. (Ami levezethető, az logikailag is következik.)
- Az axiómarendszer **teljes**, ha $\forall$ F és $\forall$ $X \rightarrow Y$ esetén

  $F \models X \rightarrow Y \;\Rightarrow\; F \vdash X \rightarrow Y$. (Minden logikai következmény levezethető.)
- $F^+ := \{X \rightarrow Y \mid F \models X \rightarrow Y\}$ F összes logikai következménye (F logikai lezártja)
- $F^* := \{X \rightarrow Y \mid F \vdash X \rightarrow Y\}$ F-ből levezethető összes függőség (F levezethetőség szerinti lezártja).
- Az axiómarendszer **helyes és teljes** (másképpen fogalmazva): $\forall$ F-re $F^+ = F^*$.

# Armstrong-axiómák

- A lezárások tulajdonságai:

1. $F \subseteq F^+$ és $F \subseteq F^*$ (reflexivitás)
2. $F \subseteq G$ esetén $F^+ \subseteq G^+$ és $F^* \subseteq G^*$ (monotonitás)
3. $F^{++} = F^+$ és $F^{**} = F^*$ (idempotencia)

- Általában lezárási operátornak hívunk egy halmazokon értelmezett műveletet, ha ezzel a 3 tulajdonsággal rendelkezik.
- **Bizonyítás** (* lezárás): 1,2 a levezethetőség definíciójából triviális. Emiatt $F^* \subseteq F^{**}$. Legyen $X \rightarrow Y \in F^{**}$. Vegyük az $X \rightarrow Y$ egy levezetését $F^*$-ből. A levezetésben szereplő $F^*$-beli függőségek helyett írjuk be azokat a levezetéseket, ahogy F-ből levezethetők. Az így felfújt levezetés $X \rightarrow Y$-nak lesz levezetése az F-ből, vagyis $X \rightarrow Y \in F^*$. **q.e.d.**
- **Bizonyítás** ($^+$ lezárás): 1,2 a logikai implikáció definíciójából triviális. Emiatt $F^+ \subseteq F^{++}$. Legyen $r \in SAT(F)$. Ekkor $r \in SAT(F^+)$. Legyen $X \rightarrow Y \in F^{++}$. Ekkor $r \in SAT(X \rightarrow Y)$. Tehát $F \models X \rightarrow Y$, azaz $X \rightarrow Y \in F^+$. **q.e.d.**

# Armstrong-axiómák

- További levezethető szabályok:

1. (Egyesítési szabály): $F \vdash X \rightarrow Y$ és $F \vdash X \rightarrow Z$ esetén $F \vdash X \rightarrow YZ$.
2. (Pszeudotranzitivitás): $F \vdash X \rightarrow Y$ és $F \vdash WY \rightarrow Z$ esetén $F \vdash XW \rightarrow Z$.
3. (Dekomponáló szabály): $F \vdash X \rightarrow Y$ és $Z \subseteq Y$ esetén $F \vdash X \rightarrow Z$.

- **Bizonyítás (1)**: Bővítési axióma miatt $F \vdash XX \rightarrow YX$ és $F \vdash YX \rightarrow YZ$, és $XX=X$, valamint a tranzitivitási axióma miatt $F \vdash X \rightarrow YZ$.
- **Bizonyítás (2)**: Bővítési axióma miatt $F \vdash XW \rightarrow YW$, és $YW=WY$, valamint a tranzitivitási axióma miatt $F \vdash XW \rightarrow Z$.
- **Bizonyítás (3):** Reflexivitási axióma miatt $F \vdash Y \rightarrow Z$, és tranzitivitási axióma miatt $F \vdash X \rightarrow Z$.
- **Következmény**: $F \vdash X \rightarrow Y \Leftrightarrow \forall Ai \in Y$ esetén $F \vdash X \rightarrow Ai$
- A következmény miatt <u>feltehető, hogy a függőségek jobb oldalai 1 attribútumból</u> állnak.

# Armstrong-axiómák

- A bal oldalak nem szedhetők szét:
  $\{AB \rightarrow C\} \not\models A \rightarrow C$.
  Ehhez elég egy ellenpélda:
  $r \in SAT(AB \rightarrow C)$, de
  $r \notin SAT(A \rightarrow C)$.

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

- $\{A \rightarrow C\} \vdash AB \rightarrow C$ viszont igaz:
  $A \subseteq AB$ miatt $\{A \rightarrow C\} \vdash AB \rightarrow A$,
  és $\{A \rightarrow C\} \vdash A \rightarrow C$, majd a tranzitivitás axiómát alkalmazzuk.
- (R,F) séma esetén legyen $X \subseteq R$.
- $X^{*(F)} := \{A \mid F \vdash X \rightarrow A\}$ az X attribútumhalmaz (levezethetőség szerinti) lezárása F-re nézve.

# Armstrong-axiómák

- **Következmény: $F \vdash X \rightarrow Y \Leftrightarrow Y \subseteq X^*$.**
- **Bizonyítás**:
  ($\Rightarrow$) $\forall A \in Y$ esetén a reflexivitás és tranzitivitás miatt $F \vdash X \rightarrow A$, azaz $Y \subseteq X^*$.
  ($\Leftarrow$) $\forall A \in Y \subseteq X^*$ esetén $F \vdash X \rightarrow A$, és az egyesítési szabály miatt $F \vdash X \rightarrow Y$. **q.e.d.**
- Az $X^*$ lezárási operátor.
- **Bizonyítás:**
  **1**. A reflexivitás miatt $F \vdash X \rightarrow X$, és a dekompozíciós szabály miatt $\forall A \in X$ esetén $F \vdash X \rightarrow A$, azaz $X \subseteq X^*$.
  **2.** Legyen $X \subseteq Y$ és $A \in X^*$. Ekkor $F \vdash X \rightarrow A$, továbbá a reflexivitás miatt $F \vdash Y \rightarrow X$, és a tranzitivitás miatt $F \vdash Y \rightarrow A$, azaz $A \in Y^*$. Tehát $X^* \subseteq Y^*$.
  **3**. Előző kettőből következik, hogy $X^* \subseteq X^{**}$. Az egyesítési szabályból következik, hogy $F \vdash X \rightarrow X^*$. Legyen $A \in X^{**}$. Ekkor $F \vdash X^* \rightarrow A$ és a tranzitivitási axiómából következik, hogy $F \vdash X \rightarrow A$, azaz $A \in X^*$. **q.e.d.**

# Armstrong-axiómák

- $X^{+(F)}:=\{A \mid F|=X\rightarrow A\}$ az X attribútumhalmaz (logikai) lezárása F-re nézve.
- Az $X^+$ is lezárási operátor.
- **Bizonyítás:**

  **1.** $\forall r\in SAT(X\rightarrow X)$, így $X\subseteq X^+$.

  **2.** Legyen $X\subseteq Y$. Vegyünk egy tetszőleges $r\in SAT(F)$ relációt. Ha r két sora megegyezik Y-n, akkor X-n is megegyeznek. Legyen $A\in X^+$, ekkor a két sor A-n is megegyezik, azaz $r\in SAT(Y\rightarrow A)$, tehát $A\in Y^+$.

  **3.** Az előző kettőből következik, hogy $X^+\subseteq X^{++}$. Vegyünk egy $r\in SAT(F)$ tetszőleges relációt. Ha r két sora megegyezik X-en, akkor megegyezik $X^+$ minden elemén, azaz $X^+$-on is. Legyen $A\in X^{++}$. Ekkor ez a két sor A-n is megegyezik. Tehát $A\in X^+$, vagyis $X^{++}\subseteq X^+$. **q.e.d.**
- *Be fogjuk látni, hogy az Armstrong-axiómarendszer helyes és teljes, azaz $\forall$ F-re $F^+=F^*$, és emiatt $\forall$ X-re $X^+= X^*$.*

# Armstrong-axiómák

- **Tétel:** Az Armstrong axiómarendszer helyes és teljes. ($\forall F$ és $\forall X \rightarrow Y$ esetén $F \vdash X \rightarrow Y \Leftrightarrow F \models X \rightarrow Y$. )
- **Bizonyítás** (helyesség). **A levezetés lépésszáma (k) szerinti <u>indukció</u>.**

  $k=1$. A levezetés $X1 \rightarrow Y1$. Ekkor vagy $X1 \rightarrow Y1 \in F$, vagy $Y1 \subseteq X1$. (A tranzitivitási és bővítési axióma alkalmazásához két, illetve egy függőséget már le kellett volna vezetni.) Mindkét esetben $F \models X1 \rightarrow Y1$.

  k hosszú levezetés esetén válasszuk szét az eseteket aszerint, hogy melyik szabályt alkalmaztuk utoljára. A levezetésben szereplő összes függőség legfeljebb k-1 hosszú levezetéssel kapható F-ből, így ezek mindegyike logikai következmény is az indukció miatt.

# Armstrong-axiómák

- ***Tétel**: Az Armstrong axiómarendszer helyes és teljes. ($\forall F$ és $\forall X \to Y$ esetén $F \vdash X \to Y \Leftrightarrow F \models X \to Y$. )*
- **Bizonyítás folytatása** (helyesség).
- - Ha a k-ik függőség F-ben szerepel, vagy $Yk \subseteq Xk$, akkor $F \models Xk \to Yk$.

  - Tegyük fel, hogy utoljára a tranzitivitást alkalmaztuk a levezetésben szereplő $X \to Z$, $Z \to Y$, függőségekre, azaz $Xk=X$, és $Yk=Y$. Vegyünk egy $r \in SAT(F)$ tetszőleges relációt. Ha r két sora megegyezik X-en, akkor Z-n is megegyeznek, mivel $X \to Z$ logikai következménye F-nek, továbbá Y-on is megegyeznek, mivel $Z \to Y$ is logikai következménye F-nek. Tehát $F \models X \to Y$, azaz $F \models Xk \to Yk$.

  - Tegyük fel, hogy utoljára a bővítést alkalmaztuk a levezetésben szereplő $X \to Y$ függőségre, azaz valamilyen Z-re $Xk=XZ$, és $Yk=YZ$. Vegyünk egy $r \in SAT(F)$ tetszőleges relációt. Ha r két sora megegyezik XZ-en, akkor X-en és Z-n is megegyeznek, és mivel $X \to Y$ indukció miatt logikai következménye F-nek, így Y-on is megegyeznek, tehát összességében YZ-en is megegyeznek, azaz $F \models XZ \to YZ$, vagyis $F \models Xk \to Yk$. **q.e.d**.

# Armstrong-axiómák

- ***Tétel:*** *Az Armstrong axiómarendszer helyes és teljes. ($\forall F$ és $\forall X \rightarrow Y$ esetén $F \vdash X \rightarrow Y$ $\Leftrightarrow$ $F \models X \rightarrow Y$. )*
- **Bizonyítás** (teljesség). Indirekten bizonyítjuk.

  Legyen F-nek logikai következménye az $X \rightarrow Y$, de ne legyen levezethető F-ből. ($F \models X \rightarrow Y$, de $F \nvdash X \rightarrow Y$.) Legyen r a következő kétsoros reláció.

| X* | R-X* |
|---|---|
| 000.....0000 | 000.....0000000 |
| 000.....0000 | 111.....1111111 |

# Armstrong-axiómák

<u>Ez valódi felbontás</u>, mert $X^*=R$ esetén $Y \subseteq X^*$ teljesülne, azaz $F \vdash X \rightarrow Y$ következne.

<u>Belátjuk, hogy $r \in SAT(F)$</u>. Legyen $V \rightarrow W \in F$.

- Ha $V \not\subseteq X^*$, akkor a két sor különbözik V-n, tehát $r \in SAT(V \rightarrow W)$.

- Ha $V \subseteq X^*$, és $W \subseteq X^*$, akkor a két sor megegyezik V-n, de W-n is, azaz $r \in SAT(V \rightarrow W)$.

- Ha $V \subseteq X^*$ , de $W \not\subseteq X^*$, akkor vegyünk egy $A \in W-X^*$ attribútumot. Ekkor $F \vdash X \rightarrow V$, $F \vdash V \rightarrow W$, $F \vdash W \rightarrow A$, azaz $F \vdash X \rightarrow A$, vagyis $A \in X^*$, de ez ellentmond az $A \in W-X^*$ választásnak, tehát ez az eset nem fordulhat elő.

# Armstrong-axiómák

Mivel $F \models X \rightarrow Y$ és $r \in SAT(F)$, ezért $r \in SAT(X \rightarrow Y)$ is teljesül.

Másrészt r mindkét sora megegyezik X*-on, és $X \subseteq X^*$ miatt X-en is, viszont $F \not\vdash X \rightarrow Y$ miatt $Y \not\subseteq X^*$, azaz az Y tartalmaz olyan attribútumot, amelyben különböző értékek szerepelnek, vagyis a két sor nem egyezik meg Y-on. Következésképpen $\mathbf{F \not\models X \rightarrow Y}$, ami ellentmondás. **q.e.d.**

# Armstrong-axiómák

- Az Armstrong-axiómarendszer "helyes és teljes" tulajdonságának következményei:
  - $F^+=F^*$
  - $X^+=X^*$
  - $\boxed{\mathbf{F \models X \rightarrow Y \Leftrightarrow F \vdash X \rightarrow Y \Leftrightarrow Y \subseteq X^*}}$
- Az implikációs probléma megoldásához elég az <u>X*-ot hatékonyan kiszámolni</u>.
- **Algoritmus (X* kiszámítására):**

  /* Iteráció, amíg X(n) változik */

  $X(0):=X$

  $X(n+1):= X(n) \cup \{A \mid \underline{Y \rightarrow Z \in F}, A \in Z, Y \subseteq X(n)\}$

  Ha $X(v+1)=X(v)$, akkor $X(v)=X^*$.
- Ha az (R,F,X) input leírási hossza k, akkor az algoritmus lépésszáma $O(k^2)$.
- A hatékonyság megfelelő könyveléssel lineárisssá, $O(k)$ lépésűvé tehető.

# Attribútumhalmaz lezárása

- **Ha $X(v+1)=X(v)$, akkor $X(v)=X^*$.**
- **Bizonyítás** ( $X(v)\subseteq X^*$):

  $X(0)=X\subseteq X^*$

  <u>indukció:</u> <u>$X(n)\subseteq X^*$</u>

  $X(n+1):= X(n) \cup \{A \mid Y\rightarrow Z\in F,\ A\in Z,\ Y\subseteq X(n)\}$

  $Y\subseteq X(n) \subseteq X^*$ miatt $F \vdash X\rightarrow Y$

  $Y\rightarrow Z\in F$ miatt $F \vdash Y\rightarrow Z$

  $A\in Z$ és a reflexivitási axióma miatt $F \vdash Z\rightarrow A$

  A *tranzitivitási axióma* miatt $F \vdash X\rightarrow A$, azaz $A\in X^*$. **q.e.d.**

# Attribútumhalmaz lezárása

- **Ha $X(v+1)=X(v)$, akkor $X(v)=X^*$.**
- **Bizonyítás ($X^* \subseteq X(v)$):**

  Indirekten tegyük fel, hogy van olyan $A \in X^*$, amelyre $A \notin X(v)$.
  Legyen r a következő kétsoros reláció:
  (Ez valódi felbontása R-nek, mert ha $X(v)=R$ lenne, akkor $X^* \subseteq X(v)=R$ teljesülne.)

| X(v) | R-X(v) |
|---|---|
| 000.....0000 | 000.....0000000 |
| 000.....0000 | 111.....1111111 |

  Belátjuk, hogy $r \in SAT(F)$.

  Legyen $V \rightarrow W \in F$.
  - Ha $V \not\subset X(v)$, akkor a két sor különbözik V-n, tehát $r \in SAT(V \rightarrow W)$.
  - Ha $V \subseteq X(v)$, és $W \subseteq X(v)$, akkor a két sor megegyezik V-n, de W-n is, azaz $r \in SAT(V \rightarrow W)$.
  - Ha $V \subseteq X(v)$, de $W \not\subset X(v)$, akkor vegyünk egy $B \in W-X(v)$ attribútumot. Ekkor $V \subseteq X(v)$, $V \rightarrow W \in F$, $B \in W$, azaz $B \in X(v+1)=X(v)$, de ez ellentmond a $B \in W-X(v)$ választásnak, tehát ez az eset nem fordulhat elő.

  $A \in X^*$ miatt $F \vdash X \rightarrow A \Leftrightarrow F \models X \rightarrow A$, és mivel $r \in SAT(F)$, ezért $r \in SAT(X \rightarrow A)$.
  $X \subseteq X(v)$, és $A \notin X(v)$ miatt $r \notin SAT(X \rightarrow A)$, ami ellentmondás. **q.e.d.**

# Attribútumhalmaz lezárása

R=ABCDEFG, {AB→C, B→G, CD→EG, BG →E}

**X=ABF**, **X*=?**

X(0):=**ABF**

X(1):=ABF∪{C,G}=ABCFG

X(2):=ABCFG ∪{C,G,E}=ABCEFG

X(3):=ABCEFG

X*=X(2)=**ABCEFG**

- R-nek K⊆R szuperkulcsa F-re nézve:
  - **F |= K→R ⇔ R⊆ K*(⊆ R ) ⇔ R=K***
- R-nek K⊆R kulcsa F-re nézve:
  - **F |= K→R ⇔ R⊆ K*(⊆ R ) ⇔ R=K***
  - minden X ⊂ K valódi részhalmazra:
    - **F |≠ X→R ⇔ R ≠ X***

# Veszteségmentes dekompozíció

- **(R,F)** sémát felbontunk **(R1,F1),...,(Rk,Fk)** részekre úgy, hogy minden R-hez tartozó r lehetséges előfordulás (azaz $r \in SAT(F)$) esetén a felbontások ugyanazt az információt tárolják mint r.
- d={(R1,F1),...,(Rk,Fk)} az (R,F) dekompozíciója, ha nem marad ki attribútum, azaz $R1 \cup ... \cup Rk = R$.
- r és (r1,...,rk) információtartalma megegyezik, ha lekérdezésekkel r-ből, illetve (r1,...,rk)-ból ugyanazokat lehet megkapni, azaz
  - minden Q relációs algebrai kifejezéshez létezik olyan Q' relációs algebrai kifejezés, amire Q(r)=Q'(r1,...,rk), és fordítva:
  - minden Q' relációs algebrai kifejezéshez létezik olyan Q relációs algebrai kifejezés, amire Q(r)=Q'(r1,...,rk).
- ri:= $\Pi_{Ri}(r)$
- Speciálisan Q(r)=r esetén ri-kből milyen Q'-vel lehet visszanyerni az r-t?
- Q'(r1,...,rk):=r1|><|...|><|rk mikor jó erre?
- **d={(R1,F1),...,(Rk,Fk)}** az **(R,F)** veszteségmentes dekompozíciója, ha minden **$r \in SAT(F)$** esetén **$r = \Pi_{R1}(r)$|><|...|><| $\Pi_{Rk}(r)$**

# Veszteségmentes dekompozíció

| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

?

=

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

|><|

| B | C |
|---|---|
| 0 | 0 |
| 0 | 1 |

=

| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |
| 0 | 0 | 1 |
| 1 | 0 | 0 |

- Új sorok keletkeztek.
- $r \subseteq \Pi_{R1}(r)|><|...|><|\Pi_{Rk}(r)$ <u>mindig teljesül</u>:
  - $t \in r$ esetén minden i-re $ti=t(Ri) \in \Pi_{Ri}(r)$, azaz
    $t = t1|><|...|><|tk \in \Pi_{R1}(r)|><|...|><|\Pi_{Rk}(r)$
- A veszteségmentességhez elég ellenőrizni, hogy minden $r \in SAT(F)$ esetén $r \supseteq \Pi_{R1}(r)|><|...|><|\Pi_{Rk}(r)$.

# Veszteségmentes dekompozíció

- Legyen $m_d(r) := \Pi_{R1}(r) |><| \ldots |><| \Pi_{Rk}(r)$.
- Kijött, hogy **minden r-re $r \subseteq m_d(r)$.**
- **Minden i-re $\Pi_{Ri}(m_d(r)) = \Pi_{Ri}(r)$**, ugyanis
  - $\Pi$ monoton, tehát $\Pi_{Ri}(m_d(r)) \supseteq \Pi_{Ri}(r)$,
  - legyen $t' \in \Pi_{Ri}(m_d(r))$, akkor van olyan $ti \in \Pi_{Ri}(r)$, hogy $t' = \Pi_{Ri}(t1 |><| \ldots |><| tk) = ti$, azaz $t' \in \Pi_{Ri}(r)$.
- **$m_d(m_d(r)) = m_d(r)$**, ugyanis
  - $m_d(m_d(r)) = \Pi_{R1}(m_d(r)) |><| \ldots |><| \Pi_{Rk}(m_d(r)) =$
    $= \Pi_{R1}(r) |><| \ldots |><| \Pi_{Rk}(r) = m_d(r)$.
- Hogy lehet (R,F,d) esetén eldönteni, hogy d veszteségmentes dekompozíciója-e R-nek F-re nézve?

# Veszteségmentes dekompozíció

- **Chase algoritmus a veszteségmentesség eldöntésére:**

  **INPUT:** (R(A1,...,An),F,d=(R1,...,Rk))
  /* feltehető, hogy F-ben a jobboldalak 1 attribútumot tartalmaznak */

  Képezünk egy r(0),r(1),r(2),...,r(v) relációkból álló sorozatot.

  **OUTPUT:** d veszteségmentes F-re $\Leftrightarrow$ $(a_1,...,a_n) \in r(v)$.

  Az r(0) kiindulási reláció ti sor j-ik eleme (i=1..k, j=1..n):

  – **t(i,j):= $a_j$**, ha **Aj$\in$ Ri,**
  – **t(i,j):= $b_{i,j}$** ,ha **Aj$\notin$ Ri.**

  (Tehát r(0)-nak annyi sora van, amennyi tagja van a d dekompozíciónak.)

# Veszteségmentes dekompozíció

A sorozatban az r(p+1) relációt úgy kapjuk, hogy az F valamelyik alkalmazható függőségét alkalmazzuk az r(p)-re, és ezt addig csináljuk, amíg lehet, azaz r(v)-re már F egyik függősége sem alkalmazható.

Egy $X \rightarrow Aj \in F$ alkalmazható r(p)-re, ha r(p) megsérti az $X \rightarrow Aj$ függőséget. Az **alkalmazás** azt jelenti, hogy értékeket azonosítunk a reláció Aj oszlopában, pontosabban

> $t,t' \in r(p)$, és $t[X]=t'[X]$, de $t(Aj) \neq t'(Aj)$, akkor a t(Aj) és t'(Aj) érték közül az **egyiket lecseréljük a másikra** a teljes Aj oszlopban.

**Preferencia** a cserénél:

- ha mindkettő $b._j$ szimbólum, akkor mindegy, hogy melyiket cseréljük a másikra,

- ha egyik $a_j$, a másik $b._j$ szimbólum, akkor a $b._j$-t cseréljük $a_j$-re.

- *az algoritmus garantálja, hogy az Aj oszlopban csak j indexű* $\boldsymbol{a_j}$ *szerepelhet, így a különböző* $a_j$*-k esete nem fordulhat elő.*

# Veszteségmentes dekompozíció

Minden cserénél eggyel csökken a relációban szereplő különböző szimbólumok száma. Új szimbólum nem jön be, így az algoritmus meg fog állni.

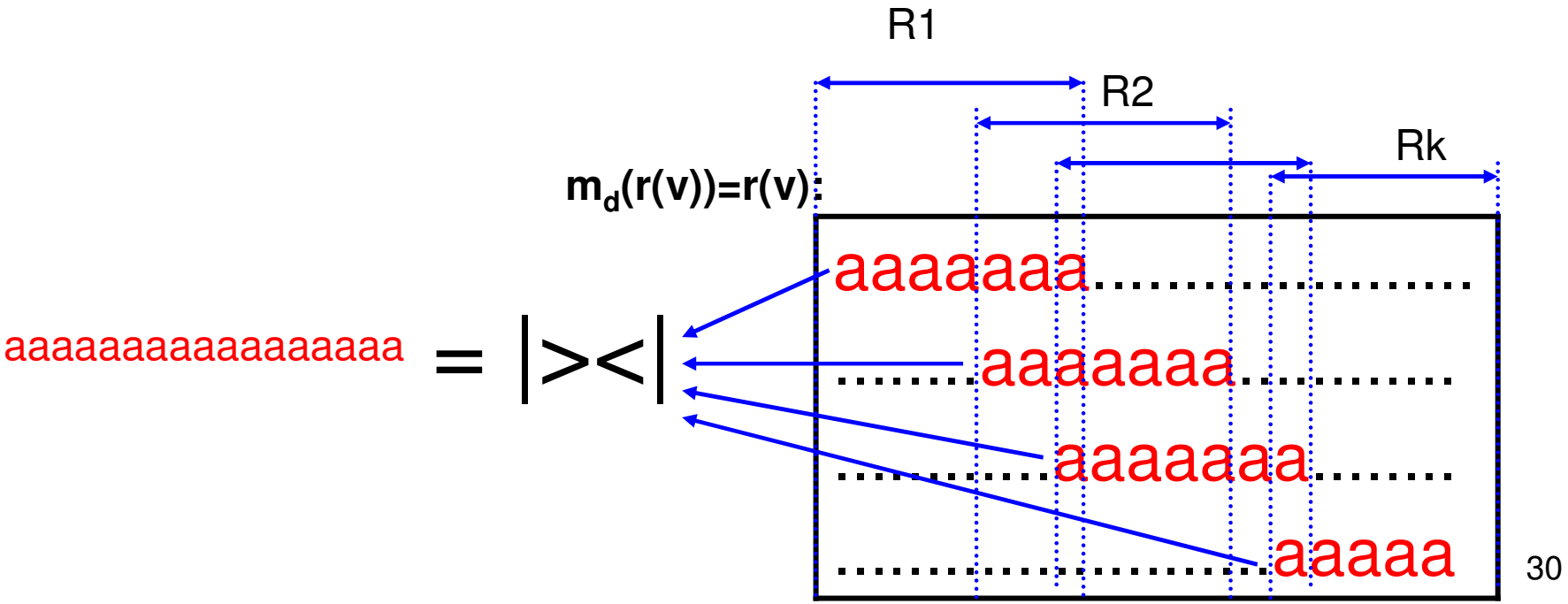
- **d veszteségmentes F-re $\Rightarrow$ $(a_1,\ldots,a_n)\in r(v)$.**
- **Bizonyítás.**
  Mivel r(v)-re már F függőségei nem alkalmazhatók, ezért $r(v)\in SAT(F)$.

  A veszteségmentesség miatt $m_d(r(v))=r(v)$.

# Veszteségmentes dekompozíció

Mivel a-t sosem cserélünk le b-re, ezért az r(0)-ban szereplő **a-k** végig a helyükön maradnak, azaz r(v)-ben is ugyanott szerepelnek.

Emiatt minden i=1..k-ra a $\Pi_{Ri}(r(v))$ vetületben lesz **csupa a-s** sor, és ezek összekapcsolása, ami pont az $(a_1,...,a_n)$ sor, benne lesz $m_d(r(v))$-ben, azaz r(v)-ben.



# Veszteségmentes dekompozíció

- **d veszteségmentes F-re $\Leftarrow$ $(a_1,\ldots,a_n)\in r(v)$.**
- **Bizonyítás.**

Az r(0),r(1),r(2),...,r(v) relációkhoz Q(0),Q(1),...,Q(v) lekérdezéseket feleltetünk meg. (A lekérdezéseket **tartománykalkulusban** adjuk meg.)

$$Q(0):= \{a_1,\ldots,a_n|\exists b..\exists b..(r(t1) \wedge\ldots\wedge r(tk))\},$$

ahol az r(0) táblában szereplő összes b..-t kvantáljuk, és ti az r(0) tábla i-edik sora (i=1..k).

Például: R=ABC, d=(AB,BC), F={B $\rightarrow$C}

r(0):

| A | B | C |
|---|---|---|
| $a_1$ | $a_2$ | $b_{1,3}$ |
| $b_{2,1}$ | $a_2$ | $a_3$ |

$$Q(0):= \{a_1,a_2,a_3|\exists b_{1,3}\exists b_{2,1}(r(a_1,a_2,b_{1,3})\wedge r(b_{2,1},a_2,a_3))\}$$

# Veszteségmentes dekompozíció

Q(0) lekérdezés értelme:

- tetszőleges ABC sémájú r relációból azokat az $a_1, a_2, a_3$ sorokat választja ki, amelyek a formulát igazzá teszik. (r(t) igaz akkor és csak akkor, ha $t \in r$.)

Belátható, hogy $Q(0)(r)=m_d(r)$.

A példában így $Q(0)(r)= \Pi_{AB}(r) |\!\!><\!\!| \Pi_{BC}(r)$.

Q(p+1)-t a Q(p)-ből úgy kapjuk, ahogy r(p+1)-et az r(p)-ből, azaz Q(p)-ben ugyanazokat a változócseréket hajtjuk végre, mint r(p)-ben, és a *lecserélt b.. változóra vonatkozó ∃b.. tagot is elhagyjuk a formulából.*

# Veszteségmentes dekompozíció

Példában B $\rightarrow$C alkalmazásával kapjuk r(0)-ból r(1)-et.

r(0)

| A | B | C |
|---|---|---|
| $a_1$ | $a_2$ | $\cancel{b_{1,3}}$ $a_3$ |
| $b_{2,1}$ | $a_2$ | $a_3$ |

r(1)

| A | B | C |
|---|---|---|
| $a_1$ | $a_2$ | $a_3$ |
| $b_{2,1}$ | $a_2$ | $a_3$ |

$Q(0):=\{a_1,a_2,a_3 | \exists b_{1,3} \exists b_{2,1} (r(a_1,a_2,b_{1,3}) \wedge r(b_{2,1},a_2,a_3))\}$

$Q(1):=\{a_1,a_2,a_3 | \quad \exists b_{2,1} (r(a_1,a_2,a_3) \wedge r(b_{2,1},a_2,a_3))\}$

# Veszteségmentes dekompozíció

Ha $Q(p+1)$-et $Q(p)$-ből az $X \rightarrow Ai$ függőség alkalmazásával kaptuk, akkor tetszőleges $r \in SAT(X \rightarrow Ai)$ esetén $Q(p+1)(r)=Q(p)(r)$.

- A példában tetszőleges $B \rightarrow C$-t kielégítő r esetén $Q(1)(r)=Q(0)(r)$.

Ha $r \in SAT(F)$, akkor így

$$\underline{m_d(r)=Q(0)(r)=Q(1)(r)=\ldots=Q(v)(r)}$$

Q(v)-nek milyen relációs algebrai kifejezés fog megfelelni?

# Veszteségmentes dekompozíció

$(a_1,\ldots,a_n)\in r(v)$ miatt:

$Q(v):= \{a_1,\ldots,a_n|\exists b..\exists b..(r(a_1,\ldots,a_n) \wedge\ldots\wedge r(\ldots))\}$,

a formula csak úgy lehet igaz, ha $a_1,\ldots,a_n\in r$, azaz tetszőleges r esetén $Q(v)(r)\subseteq r$.

Tehát $r\in SAT(F)$ esetén

$\underline{m_d(r)=Q(0)(r)=Q(1)(r)=\ldots=Q(v)(r)\subseteq r}$,

és mivel $m_d(r)\supseteq r$ mindig teljesül, így kijött, hogy $r\in SAT(F)$ esetén $\mathbf{m_d(r)=r}$, azaz <u>d veszteségmentes dekompozíció F-re nézve</u>.

**q.e.d.**

# Veszteségmentes dekompozíció

- **Speciális eset: 2 részre vágás**
- **d=(R1,R2) veszteségmentes F-re akkor és csak akkor, ha**

  $F \models R1 \cap R2 \rightarrow R1 - R2$ **vagy**

  $F \models R1 \cap R2 \rightarrow R2 - R1$.
- **Bizonyítás.**

  $F \models R1 \cap R2 \rightarrow R1 - R2 \Leftrightarrow (R1 - R2) \subseteq (R1 \cap R2)^+$

  $F \models R1 \cap R2 \rightarrow R2 - R1 \Leftrightarrow (R2 - R1) \subseteq (R1 \cap R2)^+$

  A **chase algoritmust** és a **lezárást** párhuzamosan hajtsuk végre, azaz a függőségeket ugyanolyan sorrendben alkalmazzuk mindkét algoritmusban.

# Veszteségmentes dekompozíció

r(0):

| R1∩R2 | R1 – R2 | R2 – R1 |
|---|---|---|
| aaaaaaaaaaa | aaaaaaaaaaa | bbbbbbbbbbb |
| aaaaaaaaaaa | bbbbbbbbbbb | aaaaaaaaaaa |

Y  Y→A  A

X(0)=R1∩R2, <u>X(0) oszlopaiban azonos a szimbólumok szerepelnek</u>.

*Feltehető, hogy F-ben egyeleműek a jobb oldalak.*

X(1)=X(0)∪{A|Y→A∈F, Y⊆X(0)}, így Y→A függőségeket a chase algoritmusban alkalmazva <u>X(1) minden oszlopában azonos a szimbólumok fognak szerepelni</u>.

<u>Teljes indukcióval</u> belátható, hogy egy b szimbólumot akkor és csak akkor cserélünk a-ra egy A oszlopban, ha A∈(R1∩R2)$^+$ . **q.e.d.**

# Függőségőrzés

- A dekompozíciókban érvényes függőségekből következzen az eredeti sémára kirótt összes függőség.
- Emiatt nem kell összekapcsolni a dekomponált táblákat ahhoz, hogy biztosak legyünk benne, hogy az eredetileg kirótt összes függőség teljesül.
- Milyen függőségek lesznek érvényesek a dekompozíció sémáiban?
- Függőségek vetülete:
  - Adott (R,F), és $Ri \subseteq R$ esetén:

    $\Pi_{Ri}(F) := \{ X \rightarrow Y \mid F \models X \rightarrow Y , XY \subseteq Ri \}$
- $\models$ helyett $\vdash$ is írható a vetület definíciójában.

# Függőségőrzés

- Adott (R,F) esetén d=(R1,...,Rk) **<u>függőségőrző dekompozíció</u>** akkor és csak akkor, ha

  minden $X \rightarrow Y \in F$ esetén

  $$\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F) \models X \rightarrow Y$$

- Ha $F \subseteq (\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F))^{+}$ , akkor mindkét oldalt lezárva:

  $F^{+} \subseteq (\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F))^{++} = (\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F))^{+}$

- Másrészt $\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F) \subseteq F^{+}$, mert a vetületekben F következményei közül választunk. Mindkét oldal lezártját véve: $(\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F))^{+} \subseteq F^{++} = F^{+}$, tehát a <u>függőségőrzés ekvivalens</u> azzal, hogy

  $$(\Pi_{R1}(F) \cup \ldots \cup \Pi_{Rk}(F))^{+} = F^{+}.$$

# Függőségőrzés

- *Hogy lehet a függőségőrzést ellenőrizni?*
- $\Pi_{R1}(F)\cup...\cup\Pi_{Rk}(F)\models X\rightarrow Y$ implikációt kell minden $X\rightarrow Y\in F$ függőségre ellenőrizni.
- Jelöljük G-vel a bal oldalt: $G:=\Pi_{R1}(F)\cup...\cup\Pi_{Rk}(F)$.
- **$G\models X\rightarrow Y \Leftrightarrow Y\subseteq X^{+(G)}$** miatt elő kell állítani az F függőségeiben szereplő bal oldalak **G szerinti lezárását.**
- A probléma az, hogy G számossága nagyon nagy lehet, még kis számosságú F esetén is, hiszen F összes következményei közül kell kiválasztani a vetületekbe eső függőségeket.
- ***Hogy lehet $X^{+(G)}$ –t kiszámolni anélkül, hogy G-t előállítanánk***?

# Függőségőrzés

- **$X^{+(G)}$ kiszámítása G előállítása nélkül.**
- Iteráció:

  **Z:=X**

  **while Z változik do**

  **for i:=1..k do**

  **$Z:=Z \cup ((Z \cap Ri)^{+(F)} \cap Ri)$**
- Az F szerinti lezárást **hatékonyan** tudjuk kiszámolni!
- Jelöljük $Z_f$ –fel az algoritmussal kapott attribútumhalmazt.
- **Állítás: $X^{+(G)} = Z_f$.**
- **Bizonyítás**. ($\supseteq$) Minden lépésben olyan $A \in Ri$ attribútumukat adunk Z-hez, amelyekre $F \models Z \cap Ri \rightarrow A$, azaz $Z \cap Ri \rightarrow A \in \Pi_{Ri}(F)$. Az $X^{+(G)}$-t kiszámító algoritmusban is pont ilyen függőségeket kell alkalmazni.

  ($\subseteq$) $X^{+(G)}$ kiszámítása során bármelyik új attribútum bekerülésekor G-beli, $U \rightarrow V \in \Pi_{Ri}(F)$ függőséget alkalmazunk. Indukcióval belátható, hogy $U \subseteq Z_f$, és akkor V is bekerül $Z_f$-be, azaz $X^{+(G)} \subseteq Z_f$. **q.e.d.**

# Függőségőrzés

- **R=ABCD**, **F={A→B,B→C,C→D,D→A}**, **d={AB,BC,CD}**
  $A \to B \in \Pi_{AB}(F)$, $B \to C \in \Pi_{BC}(F)$, $C \to D \in \Pi_{CD}(F)$
  Vajon megőrzi-e a $D \to A$ függőséget?
  **Z=D**
  $Z=D \cup ((D \cap AB)^+ \cap AB)=D \cup (\emptyset^+ \cap AB)=D$
  $Z=D \cup ((D \cap BC)^+ \cap BC)=D \cup (\emptyset^+ \cap BC)=D$
  $Z=D \cup ((D \cap CD)^+ \cap CD)=D \cup (D^+ \cap CD)=D \cup (ABCD \cap CD)=CD$
  $Z=CD \cup ((CD \cap AB)^+ \cap AB)=CD \cup (\emptyset^+ \cap AB)=CD$
  $Z=CD \cup ((CD \cap BC)^+ \cap BC)=CD \cup (C^+ \cap BC)=CD \cup (ABCD \cap BC)=BCD$
  $Z=BCD \cup ((BCD \cap CD)^+ \cap CD)=BCD \cup (CD^+ \cap CD)=BCD \cup (ABCD \cap CD)=BCD$
  $Z=BCD \cup ((BCD \cap AB)^+ \cap AB)=BCD \cup (B^+ \cap AB)=BCD \cup (ABCD \cap AB)=ABCD$
  Tovább már nem változik a Z. Tehát $A \in D^{+(G)}=ABCD$, vagyis ezt a **függőséget is megőrzi** a dekompozíció.
- A <u>függőségőrzésből nem következik a veszteségmentesség</u>:
  **R=ABCD, F= {A→B,C→D}**, **d={AB,CD}** függőségőrző, de nem veszteségmentes.
- A <u>veszteségmentességből nem következik a függőségőrzés:</u>
  **R=ABC, F= {AB→C,C→A}**, **d={AC,BC}** veszteségmentes, de nem függőségőrző,
  mert $C \notin AB^{+(G)}=AB$, az algoritmus alapján.

# Függőségőrzés

- Néhány további összefüggés:
- **$F1 \subseteq F2$ esetén tetszőleges V-re $V^{+(F1)} \subseteq V^{+(F2)}$.**

  **Bizonyítás**. Az F2 szerinti lezárási algoritmusban használhatjuk ugyanazokat a függőségeket, mint az F1 szerinti lezárásnál, így legalább azokat az attribútumukat megkapjuk, mint az F1 szerinti lezárás esetén. **q.e.d.**
- **$V^{+(F)} = V^{+(F^+)}$.**

  **Bizonyítás.** $F \subseteq F^+$ mindig teljesül, így tetszőleges $V^{+(F)} \subseteq V^{+(F^+)}$. Másrész tetszőleges $F^+$ halmazban szereplő $X \rightarrow Y$ függőségnek van levezetése F-ből, így a lezárási algoritmusban $X \rightarrow Y$ alkalmazása helyett vegyük a levezetésében szereplő F-beli függőségeket, és ezeket alkalmazva ugyanazt kapjuk, mintha $X \rightarrow Y$ függőséget alkalmaztuk volna. **q.e.d**.
- $G^+ \subseteq F^+$ mindig igaz, így $(Z \cap Ri)^{+(G)} = (Z \cap Ri)^{+(G^+)} \subseteq (Z \cap Ri)^{+(F^+)} = (Z \cap Ri)^{+(F)}$.

# Normálformák (BCNF)

- A nem triviális függőségek redundanciát okozhatnak.
- A <u>normálformák a redundanciát csökkentik</u>.
- *Feltesszük, hogy a lehetséges relációkat leíró F halmazban már minden függőség jobb oldala 1 attribútumot tartalmaz.*

- 1. **Definíció**. **R** relációséma **Boyce-Codd normálformában** (BCNF-ben) van az **F-re nézve**, ha tetszőleges $XA \subseteq R$, $A \notin X$ és $\boxed{\mathbf{F} \models \mathbf{X} \rightarrow \mathbf{A}}$ esetén $F \models X \rightarrow R$ (azaz X az R szuperkulcsa F-re nézve).
- 2. **Definíció**. **R** relációséma **Boyce-Codd normálformában** (BCNF-ben) van az **F-re nézve**, ha tetszőleges $XA \subseteq R$, $A \notin X$ és $\boxed{\mathbf{X} \rightarrow \mathbf{A} \in \mathbf{F}}$ esetén $F \models X \rightarrow R$ (azaz X az R szuperkulcsa F-re nézve).
- X az R szuperkulcsa F-re nézve, ha tartalmazza R-nek legalább egy F-re vonatkoztatott kulcsát. (Röviden: tartalmaz kulcsot.)

# Normálformák (BCNF)

- **A két definíció ekvivalens:**
- **Bizonyítás**. **(1.⇒2.)** Mivel $X \rightarrow A \in F$-ből következik, hogy $F \models X \rightarrow A$, így ha egy R az 1. definíció szerint BCNF-ben van, akkor a 2. definíció szerint is.

  **(1.⇐2.)** Vegyünk egy nem triviális $F \models X \rightarrow A$ függőséget. Ekkor $A \in X^+ = XA_1A_2...A...$ (lezárási algoritmussal meghatározott sorozat). $\mathbf{A \notin X}$, azaz legalább 1 lépést végrehajtottunk a lezárási algoritmusban. Eszerint van olyan nem triviális $Y \rightarrow A_1 \in F$, melyre $Y \subseteq X$, de akkor a 2. definíció miatt $F \models Y \rightarrow R$, és $F \models X \rightarrow Y$ miatt $F \models X \rightarrow R$. **q.e.d.**
- Adott **(R,F)** és d=(R1,...,Rk) esetén d az R BCNF dekompozíciója, ha minden i-re Ri **BCNF** a $\Pi_{Ri}(F)$-re nézve.

# Normálformák (BCNF)

- Vajon van-e mindig függőségőrző BCNF dekompozíció?
- **R=ABC**, F={AB→C,C→A}

  R kulcsai: AB, BC (lezárási algoritmussal ellenőrizhető). Az F függőségeinek baloldalai közül AB tartalmaz kulcsot, de C nem, azaz R nincs BCNF-ben F-re, ezért nem triviális módon dekomponálni kell BCNF sémákra. Az összes lehetséges felbontást végignézve az AB→C függőséget nem fogja megőrizni egyik dekompozíció sem. Tehát ebben a példában nem létezik függőségőrző BCNF dekompozíció.
- **Következmény**: **Tetszőleges (R,F) esetén nincs mindig függőségőrző, BCNF dekompozíció.**
- Ok: Túl erős volt a BCNF, ezért majd gyengítjük a definíciót, és így kapjuk majd a 3. normálforma definícióját.
- *Van-e minden (R,F) esetén veszteségmentes BCNF dekompozíció?* (A válasz IGEN lesz.)

# Normálformák (BCNF)

- **Ha d=(R1,...,Rk) az (R,F) veszteségmentes dekompozíciója, és d'=(S1,S2) az (R1,$\Pi_{R1}$(F)) veszteségmentes dekompozíciója, akkor d''=(S1,S2,R2,...,Rk) az (R,F) veszteségmentes dekompozíciója.**
- **Bizonyítás**. $r \in SAT(F)$ esetén $\Pi_{R1}(r) \in SAT(\Pi_{R1}(F))$, másrész $Si \subseteq R1$ miatt $\Pi_{Si}(r) = \Pi_{Si}(\Pi_{R1}(r))$, (i=1,2), így
$m_{d''}(r) = \Pi_{S1}(r) |\!\!><\!\!| \Pi_{S2}(r) |\!\!><\!\!| \Pi_{R2}(r) |\!\!><\!\!| \ldots |\!\!><\!\!| \Pi_{Rk}(r)$
$= \Pi_{S1}(\Pi_{R1}(r)) |\!\!><\!\!| \Pi_{S2}(\Pi_{R1}(r)) |\!\!><\!\!| \Pi_{R2}(r) |\!\!><\!\!| \ldots |\!\!><\!\!| \Pi_{Rk}(r) =$
$\Pi_{R1}(r) |\!\!><\!\!| \ldots |\!\!><\!\!| \Pi_{Rk}(r) = r$. **q.e.d.**

# Normálformák (BCNF)

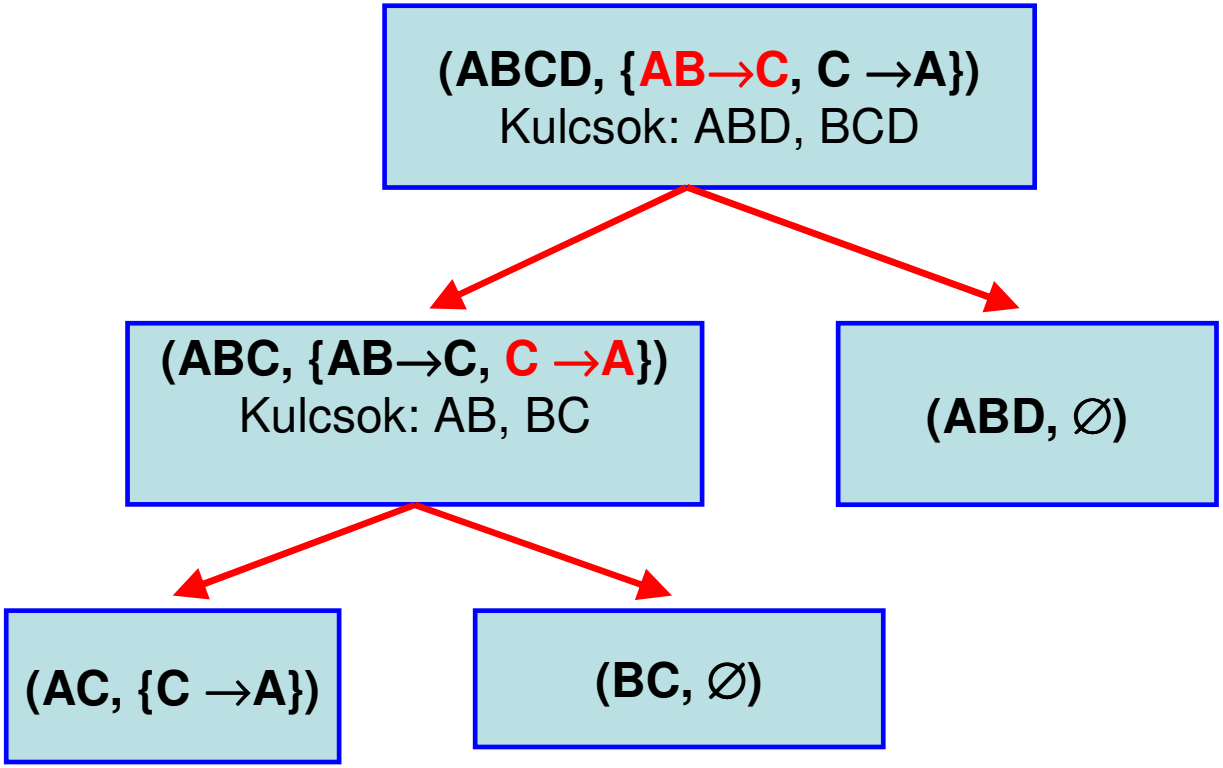
- **Minden 2 attribútumos séma tetszőleges F-re nézve BCNF-ben van.**
- **Bizonyítás**. R=AB. Milyen F függőségi halmaz jöhet szóba? **F1={A→B,B→A}, F2={A→B}, F3={B→A}.** Mindegyik esetben a bal oldalak tartalmaznak kulcsot. **q.e.d.**
- **Naiv algoritmus veszteségmentes BCNF dekompozíció előállítására:**

  Ha R BCNF-ben van, akkor megállunk, egyébként van olyan nem triviális X→A, amely R-ben teljesül, de megsérti a BCNF-et, azaz X nem szuperkulcs. Ekkor **R helyett vegyük az (XA,R-A)** dekompozíciót. XA≠R, mert különben X szuperkulcs lenne. Így *mindkét tagban csökken az attribútumok száma.*

  **XA∩(R-A)=X→A=XA-(R-A)**, azaz a kéttagú dekompozícióknál bizonyított állítás miatt *veszteségmentes kettévágást kaptunk*. A kettévágásokat addig hajtjuk végre, amíg minden tag BCNF-ben nem lesz. Legrosszabb esetben a 2 oszlopos sémákig kell szétbontani.
- Tetszőleges (R,F) esetén veszteségmentes, BCNF dekompozíciót kapunk, de nem hatékonyan, mivel nem mondtuk meg, hogy lehet X→A függőséget találni, ha nincs BCNF-ben a séma.

# Normálformák (BCNF)

**R=ABCD, F={AB→C, C →A}**




Tehát **d=(AC,BC,ABD)** veszteségmentes BCNF dekompozíció.

(∅ azt jelenti, hogy csak a triviális függőségek teljesülnek a sémában.)

# Normálformák (BCNF)

- Előfordulhat, hogy függőségőrző is a kapott dekompozíció, de ezt le kell ellenőrizni.
- Nem egyértelmű a veszteségmentes BCNF dekompozíció (más sorrendben választva a függőségeket más kettévágásokat kaphatunk).
- Hogy lehet hatékonnyá tenni a módszert?
- **Ha (R,F) nincs BCNF-ben, akkor van olyan $A,B \in R$, amelyre $F \models (R-AB) \rightarrow A$.**

  **Bizonyítás.** Van olyan $F \models X \rightarrow A$, hogy X nem szuperkulcs, ezért $XA \neq R$. Vegyünk egy $B \notin XA$ attribútumot. Ekkor $R-AB \supseteq X$, vagyis $F \models (R-AB) \rightarrow X$,

  és alkalmazzuk a tranzitivitást. **q.e.d.**
- **Következmény: Ha az összes A,B párt véve $F \not\models (R-AB) \rightarrow A$, akkor (R,F) BCNF-ben van.** (Elégséges feltétel.)
- Az összes A,B pár esetén $A \notin (R-AB)^+$ vizsgálata polinomiális, mert ($O(|R|^2)$) párra a lezárást kell képezni ($O(|R|+|F|)$).

# Normálformák (BCNF)

- **Nem szükséges feltétel:**

  R=ABC, F={C→A,C→B} kulcsa C, tehát BCNF-ben van, mégis F|=(R-AB)→A, mivel (R-AB)=C.
- A szétvágások során a vetületekben teljesülő függőségeket nem kell kiszámolni a következő állítás miatt.
- **Adott (R,F), S⊆R, XA⊆S esetén**

  $$A \in X^{+(\Pi_S(F))} \Leftrightarrow A \in X^{+(F)}.$$
- **Bizonyítás. (⇒)** $\Pi_S(F))|=X\rightarrow A$, továbbá $F|= \Pi_S(F)$ (azaz $\Pi_S(F)$ minden függősége az F következménye), így a tranzitivitás miatt $F|= X\rightarrow A$.

  **(⇐)** $F|= X\rightarrow A$ miatt $X\rightarrow A \in \Pi_S(F)$, azaz $\Pi_S(F)|=X\rightarrow A$. **q.e.d.**

# Normálformák (BCNF)

- **Adott (R,F), R2⊆R1⊆R esetén $\Pi_{R2}(\Pi_{R1}(F))=\Pi_{R2}(F)$.**
  **Bizonyítás**. Definíció miatt $\Pi_{R2}(F)= \Pi_{R2}(F^+)$, másrészt $\Pi_{R1}(F)\subseteq F^+$. Így $\Pi_{R2}(\Pi_{R1}(F))\subseteq\Pi_{R2}(F^+)= \Pi_{R2}(F)$. Másrészt $XY\subseteq R2$ esetén $XY \subseteq R1$ miatt $X\rightarrow Y\in \Pi_{R2}(F)$, $X\rightarrow Y\in \Pi_{R1}(F)$, így $X\rightarrow Y\in \Pi_{R2}(\Pi_{R1}(F))$. **q.e.d.**

- **Hatékony algoritmus veszteségmentes BCNF dekompozíció előállítására:**
- Input:(R,F), Output: d veszteségmentes BNCF dekompozíció
- **Z:=R**
  **repeat**
  **dekomponáljuk Z-t (Z-A,XA) részekre úgy, hogy XA BCNF a $\Pi_{XA}(F)$-re nézve, és F|=X→A.**
  **Legyen XA∈d**
  **Z:=Z-A**
  **until Z-t nem lehet tovább dekomponálni**
  **Legyen Z∈d**
- Ha Z-ben nincs olyan A,B, amelyre **F|=(Z-AB)→A**, akkor Z BCNF-ben van és tovább nem dekomponálható.
  Különben legyen A,B olyan Z-beli attribútum, amelyre **F|=(Z-AB)→A**
  Y:=Z-B
  while van olyan **A,B∈Y**, amelyre **F|=(Y-AB)→A**
  Y:=Y-B
  Return (Z-A,Y) /* A a while ciklusban utoljára beállított A attribútum, X:=Y-A */

# Normálformák (BCNF)

- **Órarend adatbázis**: Kurzus(K), Oktató(O), Időpont(I), Terem(T), Diák(D), Jegy(J)
- **Feltételek:**

  Egy kurzust csak egy oktató tarthat: $K \rightarrow O$.

  Egy helyen, egy időben csak egy kurzus lehet: $IT \rightarrow K$.

  Egy időben egy tanár csak egy helyen lehet: $IO \rightarrow T$.

  Egy diák egy tárgyból csak egy jegyet kaphat: $KD \rightarrow J$.

  Egy időben egy diák csak egy helyen lehet:

  $ID \rightarrow T$.
- R=KOITDJ $F = \{K \rightarrow O, IT \rightarrow K, IO \rightarrow T, KD \rightarrow J, ID \rightarrow T\}$

# Normálformák (BCNF)

- **R=KOITDJ F= {K→O, IT→K, IO→T, KD→J, ID→T }**
- **Z:=KOITDJ**
- **AB:=KO, mert F|=ITDJ→K**

  **Y:=KITDJ**

  **AB:=TK, mert F|=IDJ→T**

  **Y:=ITDJ**

  **AB:=TJ, mert F|=ID→T**

  **Y:=ITD tovább nem dekomponálható (BCNF)**
- **Z:=KOITDJ-T=KOIDJ**
- **AB:=OI, mert F|=KDJ→O**

  **Y:= KODJ**

  **AB:=OD, mert F|=KJ→O**

  **Y:= KOJ**

  **AB:=OJ, mert F|=K→O**

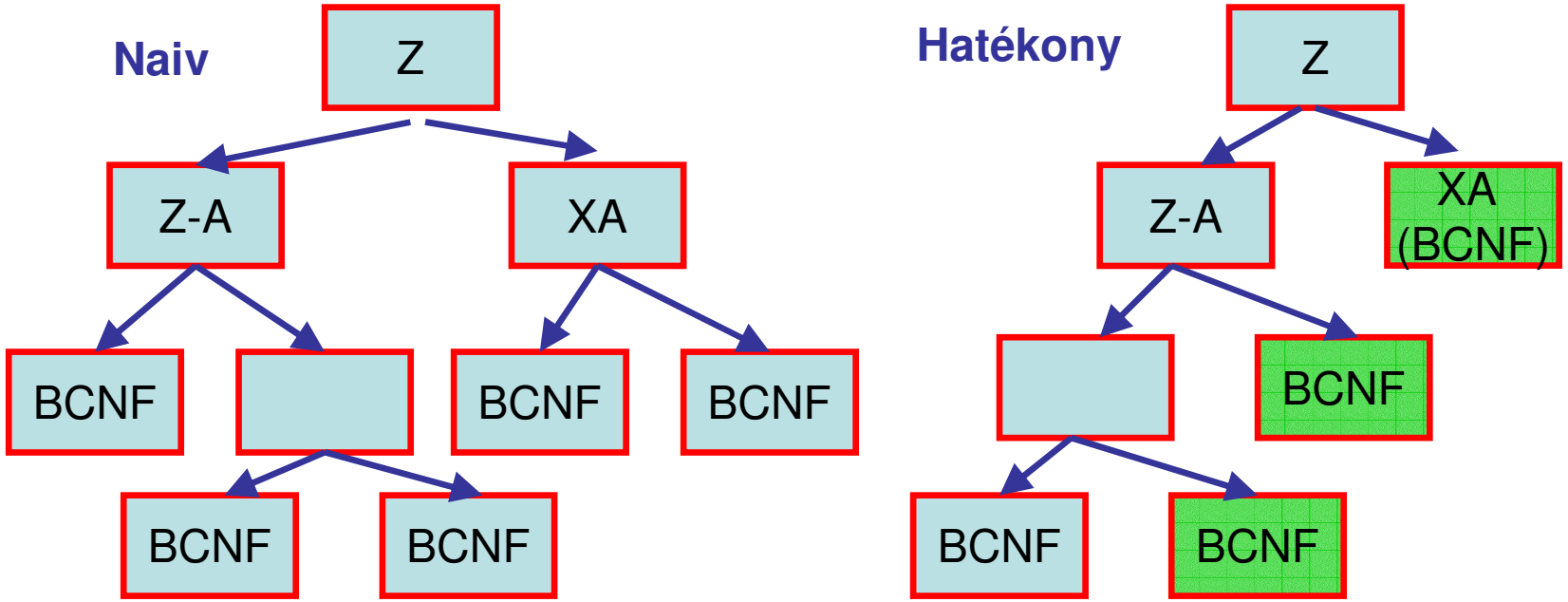
  **Y:=KO tovább nem dekomponálható (BCNF)**
- **Z:=KOIDJ-O:=KIDJ**

  **AB:=JI, mert F|=KD→J**

  **Y:=KDJ tovább nem dekomponálható (BCNF)**
- **Z:=KIDJ-J=KID tovább nem dekomponálható (BCNF)**
- Output: **d=(ITD,KO,KDJ,KID)** az (R,F) veszteségmentes BNCF dekompozíciója.

# Normálformák (BCNF)

- Az algoritmus a párok keresése, és a lezárások kiszámítása miatt polinomiális lépés számú.
- Szemben a naiv módszerrel a kettévágások során az egyik tag mindig BCNF-ben van.



# Normálformák (3NF)

- Adott (R,F) esetén $A \in R$ az R **elsődleges attribútuma** F-re nézve, ha A szerepel az R valamelyik F-re vonatkoztatott kulcsában.
- 1. **Definíció.** R relációséma **3. normálformában** (3NF-ben) van az F-re nézve, ha tetszőleges $XA \subseteq R$, $A \notin X$ és $\boxed{F \models X \rightarrow A}$ esetén $F \models X \rightarrow R$ (azaz X az R szuperkulcsa F-re nézve) vagy A az R elsődleges attribútuma F-re nézve.
- 2. **Definíció**. R relációséma **3. normálformában** (3NF-ben) van az F-re nézve, ha tetszőleges $XA \subseteq R$, $A \notin X$ és $\boxed{X \rightarrow A \in F}$ esetén $F \models X \rightarrow R$ (azaz X az R szuperkulcsa F-re nézve) vagy A az R elsődleges attribútuma F-re nézve.

# Normálformák (3NF)

- **A két definíció ekvivalens:**
- **Bizonyítás. (1.⇒2.)** Mivel $X \to A \in F$-ből következik, hogy $F \models X \to A$, így ha egy R az 1. definíció szerint 3NF-ben van, akkor a 2. definíció szerint is.

  **(1.⇐2.)** Vegyünk egy nem triviális $F \models X \to A$ függőséget. Ekkor $\mathbf{A \in X^+ = XA_1A_2 \ldots A \ldots}$ (lezárási algoritmussal meghatározott sorozat). $\mathbf{A \notin X}$, azaz <u>legalább 1 lépést végrehajtottunk</u> a lezárási algoritmusban. Eszerint van olyan nem triviális $\mathbf{Y \to A_1 \in F}$, melyre $\mathbf{Y \subseteq X}$, de akkor a 2. definíció miatt vagy $F \models Y \to R$, és ekkor $F \models X \to Y$ miatt $\mathbf{F \models X \to R}$ vagy $A_1$ elsődleges. Nézzük meg, hogy <u>$A_2$ milyen függőség miatt jön be</u> a sorba: $Y' \to A_2 \in F$ nem triviális függőség, melyre $Y' \subseteq XA_1$. Ekkor a 2. definíció miatt vagy $F \models Y' \to R$, és ekkor $F \models XA_1 \to Y'$ valamint $F \models X \to A_1$ miatt $\mathbf{F \models X \to R}$ vagy $A_2$ elsődleges. Tovább folytatva **indukcióval** belátható, hogy mivel **A** szerepel valamilyen $A_i$ tagként a sorban, ezért vagy $\mathbf{F \models X \to R}$ (azaz X szuperkulcs) vagy A elsődleges. **q.e.d.**

# Normálformák (3NF)

- Adott **(R,F)** és **d=(R1,...,Rk)** esetén d az R 3NF dekompozíciója, ha minden i-re Ri 3NF a $\Pi_{Ri}$(F)-re nézve.
- R=ABC, F={AB→C,C→A}

  **R kulcsai:** AB, BC

  Láttuk, hogy nincs BCNF-ben, és nincs függőségőrző BCNF dekompozíciója.
- **R elsődleges attribútumai**: A,B,C, így R 3NF-ben van F-re nézve, tehát nem kell tovább dekomponálni.
- Mivel a definíciót gyengítettük, így **ha (R,F) BCNF, akkor (R,F) 3NF is**.

# Normálformák (3NF)

- **Mindig létezik veszteségmentes 3NF dekompozíció, ami még függőségőrző is. (Bizonyítás később.)**
- Előtte nézzük meg, hogy lehet minimálisra csökkenteni egy F függőségi halmazt.
- **G az F minimális fedése**, ha

  - **$F^+=G^+$ (fedés),**

  - G minden függőségében a jobb oldalak egyeleműek, (**jobb oldalak minimálisak)**

  - G-ből nem hagyható el függőség, hogy F fedése maradjon, **(minimális halmaz)**

  - G függőségeinek bal oldala nem csökkenthető, hogy F fedése maradjon **(bal oldalak minimálisak).**
- Kérdések:
  - Minden F-nek létezik minimális fedése?
  - Egyértelmű-e a minimális fedés?
  - Hogyan lehet adott F esetén meghatározni egy minimális fedést, lehetőleg hatékonyan?

# Normálformák (3NF)

- Hogy lehet általában ellenőrizni, hogy teljesül-e a $F^+=G^+$ ?
- **$F^+=G^+$ akkor és csak akkor, ha $F\subseteq G^+$ és $F^+\supseteq G$.**
- **Bizonyítás**. **($\Rightarrow$)** $F\subseteq F^+=G^+\supseteq G$.

  **($\Leftarrow$)** $F\subseteq G^+$ és $F^+\supseteq G$ mindkét oldal lezárását véve: $F^+\subseteq G^{++}=G^+$ és $F^+=F^{++}\supseteq G^+$. **q.e.d.**
- Ha $F=\{X1\rightarrow Y1,...,Xk\rightarrow Yk\}$ akkor $F\subseteq G^+$ akkor és csak akkor, ha $G|=Xi\rightarrow Yi$, (i=1..k), azaz $Yi\subseteq Xi^{+(G)}$, (i=1..k).
- $F^+\supseteq G$ ellenőrzése $G=\{U1\rightarrow V1,...,Um\rightarrow Vm\}$ esetén akkor és csak akkor, ha $Vi\subseteq Ui^{+(F)}$, (i=1..m).
- Ha $G\subseteq F$, akkor $G\subseteq G^+\subseteq F^+$ miatt, $F^+=G^+$ akkor és csak akkor, ha $F\subseteq G^+$, azaz $Yi\subseteq Xi^{+(G)}$, (i=1..k). Mivel $G\subseteq F$, ezért valójában elég csak az **F-G** függőségeit vizsgálni.
- **Következmény: $G:=F-\{X\rightarrow Y\}$ esetén $F^+=G^+$ akkor és csak akkor, ha $Y\subseteq X^{+(G)}$.**

# Normálformák (3NF)

- **Mohó algoritmussal mindig található minimális fedés:**

1. **Jobb oldalak minimalizálása:**

   $X \rightarrow A1,\ldots,Ak$ függőséget cseréljük le az

   **$X \rightarrow A1, \ldots, X \rightarrow Ak$** k darab függőségre.

2. **A halmaz minimalizálása:**

   Hagyjuk el az olyan $X \rightarrow A$ függőségeket, amelyek a fedést nem befolyásolják, azaz

   while F változik

   **if $(F-\{X \rightarrow A\})^+ = F^+$ then $F := F-\{X \rightarrow A\}$;**

3. **Bal oldalak minimalizálása:**

   Hagyjuk el a bal oldalakból azokat az attribútumokat, amelyek a fedést nem befolyásolják, azaz

   while F változik

   for all $X \rightarrow A \in F$

   for all $B \in X$

   **if $((F-\{X \rightarrow A\}) \cup \{(X-\{B\}) \rightarrow A\})^+ = F^+$ then $F := (F-\{X \rightarrow A\}) \cup \{X-\{B\}) \rightarrow A\}$**

- Belátható, hogy a 3. lépés nem rontja el a halmaz minimalizálást, így minimális fedést kapunk.

# Normálformák (3NF)

- **Az algoritmusban különböző sorrendben választva a függőségeket, illetve attribútumokat, különböző minimális fedést kaphatunk.**
- **$F=\{A\rightarrow B, B\rightarrow A, B\rightarrow C, A\rightarrow C, C\rightarrow A\}$**

  $(F-\{B\rightarrow A\})^+=F^+$, mivel $F-\{B\rightarrow A\}\models B\rightarrow A$

  $F:=F-\{B\rightarrow A\}$

  $(F-\{A\rightarrow C\})^+=F^+$, mivel $F-\{A\rightarrow C\}\models A\rightarrow C$

  $F:=F-\{A\rightarrow C\}=$ **$\{A\rightarrow B,B\rightarrow C,C\rightarrow A\}$ minimális fedés,** mert több függőség és attribútum már nem hagyható el.
- **$F=\{A\rightarrow B, B\rightarrow A, B\rightarrow C, A\rightarrow C, C\rightarrow A\}$**

  $(F-\{B\rightarrow C\})^+=F^+$, mivel $F-\{B\rightarrow C\}\models B\rightarrow C$

  $F:=F-\{B\rightarrow C\}=$**$\{A\rightarrow B,B\rightarrow A,A\rightarrow C,C\rightarrow A\}$ is minimális fedés**, mert több függőség és attribútum már nem hagyható el.
- **$F=\{AB\rightarrow C, A\rightarrow B, B\rightarrow A\}$**

  $(F-\{AB\rightarrow C\}\cup\{A\rightarrow C\})^+=F^+$, mivel $(F-\{AB\rightarrow C\})\cup\{A\rightarrow C\}\models AB\rightarrow C$ és $F\models A\rightarrow C$.

  $F:=(F-\{AB\rightarrow C\}\cup\{A\rightarrow C\})=$ **$\{A\rightarrow C,A\rightarrow B,B\rightarrow A\}$ minimális fedés**, mert több függőség és attribútum már nem hagyható el.
- **$F=\{AB\rightarrow C, A\rightarrow B, B\rightarrow A\}$**

  $(F-\{AB\rightarrow C\}\cup\{B\rightarrow C\})^+=F^+$, mivel $(F-\{AB\rightarrow C\})\cup\{B\rightarrow C\}\models AB\rightarrow C$ és $F\models B\rightarrow C$.

  $F:=(F-\{AB\rightarrow C\}\cup\{B\rightarrow C\})=$ **$\{B\rightarrow C,A\rightarrow B,B\rightarrow A\}$ is minimális fedés**, mert több függőség és attribútum már nem hagyható el.

- **Algoritmus függőségőrző 3NF dekompozíció előállítására:**
- Input: (R,F)
  - Legyen **G**:={X→A,X→B,...,Y→C,Y→D,....} az **F minimális fedése.**
  - Legyen S az R sémának G-ben nem szereplő attribútumai.
  - Ha van olyan függőség G-ben, amely R összes attribútumát tartalmazza, akkor legyen d:={R}, különben legyen

**d:={S,XA, XB,...,YC,YD,...}.**

# Normálformák (3NF)

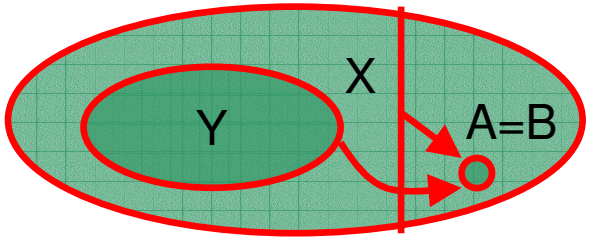
- **A d dekompozíció 3NF és függőségőrző.**
- **Bizonyítás**. Mivel G összes függősége szerepel a d valamelyik sémájában, és G az F fedése, így a függőségőrzés teljesül.

  S sémában csak a triviális függőségek teljesülnek, így S 3NF-ben van.

  XA miért van 3NF-ben? **Indirekten** tegyük fel, hogy $G \models Y \rightarrow B$ megsérti a 3NF-et (**Y nem tartalmaz kulcsot**, és **B nem elsődleges**).

  **1. eset: A=B**.

**XA:**




$Y \subseteq X$, de $\mathbf{X \neq Y}$, (mert X=Y esetén $Y=X \rightarrow A=B$ miatt **Y szuperkulcsa lenne XA-nak**).

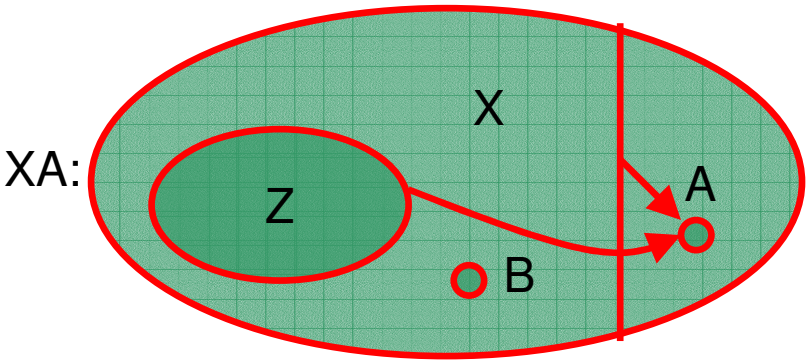
$Y \rightarrow A \models X \rightarrow A$ miatt $\mathbf{((G-\{X \rightarrow A\}) \cup \{(Y \rightarrow A\})^+ = G^+}$, azaz G-ben az $\mathbf{X \rightarrow A}$ függőség **nem lenne minimális bal oldalú**.

- **2. eset: $A \neq B$.**

  Mivel $G \models X \rightarrow A$, ezért X szuperkulcs XA-ban.

  Legyen Z egy kulcs X-ben XA-ra nézve. **B nem elsődleges**, így $B \notin Z$, azaz $Z \subseteq X$, de $Z \neq X$. Mivel $G \models Z \rightarrow XA$ és $Z \rightarrow A \models X \rightarrow A$,

  így $((G-\{X \rightarrow A\}) \cup \{(Z \rightarrow A\})^+ = G^+$, azaz G-ben az $X \rightarrow A$ függőség **nem lenne minimális bal oldalú**. **q.e.d**.

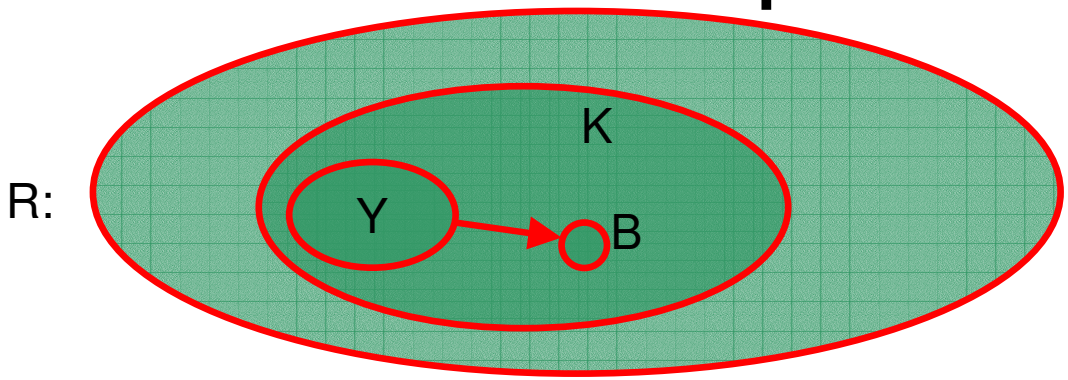


- **Algoritmus függőségőrző és veszteségmentes 3NF dekompozíció előállítására:**
- **Input: (R,F)**
  - Legyen $G:=\{X\rightarrow A, X\rightarrow B, \ldots, Y\rightarrow C, Y\rightarrow D, \ldots\}$ az **F minimális fedése.**
  - Legyen S az R sémának G-ben nem szereplő attribútumai.
  - Ha van olyan függőség G-ben, amely R összes attribútumát tartalmazza, akkor legyen $d:=\{R\}$, különben legyen K az R egy kulcsa, és legyen

$$d:=\{K, S, XA, XB, \ldots, YC, YD, \ldots\}.$$

- **A d dekompozíció 3NF és függőségőrző és veszteségmentes.**
- **Bizonyítás**. A függőségőrzés nem sérül, mivel továbbra is G minden függősége szerepel valamelyik vetületben.

  K miért van 3NF-ben? **Indirekten** tegyük fel, hogy $G|=Y\rightarrow B$ teljesül K-ban, ahol **Y nem szuperkulcsa K-nak** és **B nem szerepel K kulcsában**. Mivel $Y\subseteq K-\{B\}$, ezért $G|=K-\{B\}\rightarrow B$ és $G|=K-\{B\}\rightarrow K-\{B\}$, azaz $G|=K-\{B\}\rightarrow K$. Mivel K kulcs, ezért $G|=K\rightarrow R$, de így $G|=K-\{B\}\rightarrow R$ is teljesülne, azaz K **nem lenne minimális szuperkulcs.**



# Normálformák (3NF)

Veszteségmentesség ellenőrzése a Chase algoritmussal történik. **$K^{+(G)}$=KA1,...,Ak=R**. A G-beli függőségeket ugyanolyan sorrendben alkalmazzuk a **chase algoritmusban** és a **lezárási algoritmusban**.

**Indukcióval látható**, hogy a K-nak megfelelő sorban az Ai oszlopban a chase algoritmus végén "**a**" szimbólum fog szerepelni:

Ha **KA1,...,Ai-1$\supseteq$Y**$\rightarrow$Ai$\in$G hozza be az Ai-t, akkor **indukció miatt** a **K** és **YAi** sémához tartozó sorok az Y oszlopokban egyformán "**a**" szimbólumokat tartalmaznak, így a K sorában az Ai oszlopban is "**a**"-ra kell cserélni az értéket. K sorában csupa "**a**" szimbólum fog szerepelni, azaz veszteségmentes a dekompozíció. **q.e.d.**

| | | K | Ai |
|---|---|---|---|
| | | | |
| K | aaaaaaa | aaaaaaaaaaaaaaaaaaaa | ~~b~~ a |
| YAi | aaaaaaa | aaaaaaa | a |
| | | | |

- **Algoritmus függőségőrző és veszteségmentes 3NF redukált (kevesebb tagból álló) dekompozíció előállítására:**
- **Input: (R,F)**
  - Legyen $G:=\{X\rightarrow A, X\rightarrow B, \ldots, Y\rightarrow C, Y\rightarrow D, \ldots\}$ az **F minimális fedése**.
  - Legyen S az R sémának G-ben nem szereplő attribútumai.
  - Ha van olyan függőség G-ben, amely R összes attribútumát tartalmazza, akkor legyen $d=\{R\}$, különben legyen K az R egy kulcsa, és legyen

**$d:=\{K,S,XAB\ldots,\ldots,YCD\ldots,\ldots\}$.**

# Normálformák (3NF)

- **A d dekompozíció 3NF és függőségőrző és veszteségmentes.**
- **Bizonyítás**. A függőségőrzés nem változik, mivel $\Pi_{XA}(F) \cup \Pi_{XB}(F) \subseteq \Pi_{XAB}(F)$. A veszteségmentesség bizonyítása ugyanúgy megy, mint előbb. Az is belátható, hogy XAB is 3NF-ben van, ha XA és XB 3NF-ben van. **q.e.d**.
- Ha XAB...,YCD... tagok valamelyike szuperkulcs, akkor nem kell külön K kulcsot bevenni a dekompozícióba.
- Ha **S=** $\emptyset$, akkor S-t nem kell szerepeltetni a dekompozícióban.

# Normálformák (3NF)

- **Órarend adatbázis**: Kurzus(K), Oktató(O), Időpont(I), Terem(T), Diák(D), Jegy(J)
- **Feltételek:**

  Egy kurzust csak egy oktató tarthat: $K \rightarrow O$.

  Egy helyen, egy időben csak egy kurzus lehet: $IT \rightarrow K$.

  Egy időben egy tanár csak egy helyen lehet: $IO \rightarrow T$.

  Egy diák egy tárgyból csak egy jegyet kaphat: $KD \rightarrow J$.

  Egy időben egy diák csak egy helyen lehet: $ID \rightarrow T$.
- R=KOITDJ F= $\{K \rightarrow O, IT \rightarrow K, IO \rightarrow T, KD \rightarrow J, ID \rightarrow T\}$
- F-nek F minimális fedése, ID kulcs, és $ID \subseteq IDT$, így

  **d={KO,ITK,IOT,KDJ,IDT} 3NF függőségőrző, veszteségmentes dekompozíció**

# Többértékű függőségek

- **Dolgozó adatbázis**: Név(N), Diploma(D), Telefon(T)
- R=NDT

| N | D | T |
|---|---|---|
| Kovács | {programozó, közgazdász} | {1234567, 7654321, 1212123} |
| Szabó | {programozó, jogász} | {1234123, 1234512} |

0. normálforma: az értékek halmazok is lehetnek.

# Többértékű függőségek

- Átírás 1. normálfomára (az értékek atomi értékek)

| N | D | T |
|---|---|---|
| Kovács | {programozó, közgazdász} | {1234567, 7654321, 1212123} |
| Szabó | {programozó, jogász} | {1234123, 1234512} |

| N | D | T |
|---|---|---|
| Kovács | programozó | 1234567 |
| Kovács | programozó | 7654321 |
| Kovács | programozó | 1212123 |
| Kovács | közgazdász | 1234567 |
| Kovács | közgazdász | 7654321 |
| Kovács | közgazdász | 1212123 |
| Szabó | programozó | 1234123 |
| Szabó | programozó | 1234512 |
| Szabó | jogász | 1234123 |
| Szabó | jogász | 1234512 |

30 értéket tárolunk (redundancia)!

- Adott névhez diplomák halmaza és telefonszámok halmaza tartozik, egymástól függetlenül.
- **Név →→ Diploma**
- **Név →→ Telefon**

# Többértékű függőségek

- Dekomponáljuk 2 táblára veszteségmentesen:

| N | D |
|---|---|
| Kovács | programozó |
| Kovács | közgazdász |
| Szabó | programozó |
| Szabó | jogász |

| N | T |
|---|---|
| Kovács | 1234567 |
| Kovács | 7654321 |
| Kovács | 1212123 |
| Szabó | 1234123 |
| Szabó | 1234512 |

18 értéket tárolunk (csökkent a redundancia)

• A 2 tábla összekapcsolása visszaadná az eredeti (redundáns) táblát, vagyis veszteségmentes lenne a dekompozíció.

• A funkcionális függőség speciális többértékű függőség, például Név → Telefon esetén 1 elemű halmaz ( 1 telefonszám) tartozik minden névhez, azaz Név →→ Telefon.

# Többértékű függőségek

- **Definíció**: $X,Y \subseteq R$, $Z:=R-XY$ esetén $X \rightarrow\rightarrow Y$ **többértékű függőség. (tf)**
- **Definíció(1)**: Egy R sémájú **r reláció kielégíti** az $X \rightarrow\rightarrow Y$ függőséget, azaz $r \in SAT(X \rightarrow\rightarrow Y)$, ha $t,s \in r$ és $t[X]=s[X]$ esetén <u>létezik olyan $u,v \in r$,</u> amelyre $u[X]=v[X]=t[X]=s[X]$, $u[Y]=t[Y]$, $u[Z]=s[Z]$, $v[Y]=s[Y]$, $v[Z]=t[Z]$.

| | X | Y | Z |
|---|---|---|---|
| t | x | y1 | z1 |
| s | x | y2 | z2 |
| ∃ **u** | x | y1 | z2 |
| ∃ **v** | x | y2 | z1 |

# Többértékű függőségek

- Elég az u,v közül csak az egyik létezését megkövetelni.
- **Definíció(2)**: Egy R sémájú r reláció kielégíti az $X \rightarrow\rightarrow Y$ függőséget, azaz $r \in SAT(X \rightarrow\rightarrow Y)$, ha $t,s \in r$ és $t[X]=s[X]$ esetén <u>létezik olyan $u \in r$,</u> amelyre $u[X]=t[X]=s[X]$, $u[Y]=t[Y]$, $u[Z]=s[Z]$.
- Definíció(1)-ből következik a Definíció(2), de visszafele is igaz, mert a t,s-hez létezik u, és s,t-hez is létezik egy sor, ami pont v.

| | X | Y | Z |
|---|---|---|---|
| t | x | y1 | z1 |
| s | x | y2 | z2 |
| ∃ u | x | y1 | z2 |

# Többértékű függőségek

- Hasonló utat járunk be, mint a funkcionális függőségek esetén:
  - implikációs probléma
  - axiomatizálás
  - levezethető függőségek hatékony meghatározása (lezárás helyett függőségi bázis)
  - veszteségmentes dekompozíció
  - 4. normálforma
  - veszteségmentes 4NF dekompozíció előállítása
- Mivel kijön majd, hogy minden 4NF egyben BCNF is, amire nincs egyszerre függőségőrző és veszteségmentes dekompozíció, így 4NF-re sincs mindig.

# Többértékű függőségek

- **Jelölés:**
  - **F** funkcionális függőségek halmaza
  - **M** többértékű függőségek halmaza
  - **D** vegyes függőségek (funkcionális és többértékű függőségek) halmaza
- **SAT(függőségi halmaz)**:= **összes** olyan **reláció**, amely a függőségi halmaz minden elemét kielégíti.
- **Implikációs probléma:**
  - $F \models X \rightarrow Y \iff SAT(F) \subseteq SAT(X \rightarrow Y)$
  - $M \models X \rightarrow\rightarrow Y \iff SAT(M) \subseteq SAT(X \rightarrow\rightarrow Y)$
  - $D \models X \rightarrow Y \iff SAT(D) \subseteq SAT(X \rightarrow Y)$
  - $D \models X \rightarrow\rightarrow Y \iff SAT(D) \subseteq SAT(X \rightarrow\rightarrow Y)$

# Többértékű függőségek

- Axiomatizálás

| **Funkcionális függőségek** | **Többértékű függőségek** | **Vegyes függőségek** |
|---|---|---|
| **A1 (reflexivitás):** $Y \subseteq X$ esetén $X \rightarrow Y$. | **A4 (komplementer):** $X \rightarrow\rightarrow Y$ és $Z=R-XY$ esetén $X \rightarrow\rightarrow Z$. | **A7 (funkcionálisból többértékű):** $X \rightarrow Y$ esetén $X \rightarrow\rightarrow Y$. |
| **A2 (tranzitivitás):** $X \rightarrow Y$ és $Y \rightarrow Z$ esetén $X \rightarrow Z$. | **A5 (tranzivitás):** $X \rightarrow\rightarrow Y$ és $Y \rightarrow\rightarrow S$ esetén $X \rightarrow\rightarrow S-Y$. | **A8 (többértékűből és funcionálisból funkcionális):** $X \rightarrow\rightarrow Y$ és $W \rightarrow S$, ahol $S \subseteq Y$, $W \cap Y = \emptyset$ esetén $X \rightarrow S$. |
| **A3 (bővíthetőség):** $X \rightarrow Y$ és tetszőleges $Z$ esetén $XZ \rightarrow YZ$. | **A6 (bővíthetőség):** $X \rightarrow\rightarrow Y$ és tetszőleges $V \subseteq W$ esetén $XW \rightarrow\rightarrow YV$. | |

# Többértékű függőségek

- **Tétel** (**helyes** és **teljes** axiómarendszerek):
  - **A1,A2,A3 helyes és teljes a funkcionális függőségekre,**
  - **A4,A5,A6 helyes és teljes a többértékű függőségekre,**
  - **A1,A2,A3,A4,A5,A6,A7,A8 helyes és teljes a vegyes függőségekre.**

**Bizonyítás (részlet):** A1,A2,A3 az Armstrong-axiómarendszer, amiről már beláttuk, hogy helyes és teljes.

A másik két axiómarendszer esetén a **teljesség** bizonyítása **indirekten** történik, egy valamivel bonyolultabb ellenpélda konstruálásával, mint a funkcionális függőségek esetében.

A **helyesség** igazolása a levezetés lépésszáma szerinti **indukcióval**, és az utolsó szabály szerinti esetszétválasztással történik. **Például,** ha az utolsó függőséget **A5** alkalmazásával kaptuk: Azt kell belátni, hogy ha r kielégíti az $X \rightarrow\rightarrow Y$ és $Y \rightarrow\rightarrow S$ függőségeket, akkor kielégíti az $X \rightarrow\rightarrow S-Y$ függőséget is.

Megmutatjuk, hogy a függőség definíciójában szereplő t-hez és s-hez létezik megfelelő **v** sor. Az első függőség miatt létezik egy u sor, majd a t és u sorból a második függőség miatt létezik a **v** sor.

# Többértékű függőségek

- Az egyszerűség kedvéért tegyük fel, hogy X-nek nincs közös része se Y-nal, se S-sel. (Ha ez nem teljesül, akkor is igaz, csak akkor az ábra nem áttekinthető.) Legyen Z=R-XYS.

Y S

| | X | Y-S | S∩Y | S-Y | Z |
|---|---|---|---|---|---|
| t | x | a | b | c | d |
| s | x | e | f | g | h |
| X→→Y miatt ∃u | x | a | b | g | h |
| Y→→S miatt ∃v | x | a | b | g | d |

# Többértékű függőségek

Nézzünk meg egy másik esetet is. **Például,** ha az utolsó függőséget az **A8** alkalmazásával kaptuk: Azt kell belátni, hogy ha r kielégíti az $X \rightarrow\rightarrow Y$ és $W \rightarrow S$ függőségeket, ahol $S \subseteq Y$, $W \cap Y = \emptyset$, akkor kielégíti az $X \rightarrow S$ függőséget is.

Megmutatjuk, hogy ha r táblában t és s sor megegyezik X-en, akkor megegyezik S-en is.

A többértékű függőség miatt t és s-hez létezik egy megfelelő **u** sor, és mivel t és u a W-n megegyezik, így a funkcionális függőség miatt S-en is megegyeznek.

# Többértékű függőségek

- Az egyszerűség kedvéért megint tegyük fel, hogy X-nek nincs közös része se Y-nal, se W-vel. (Ha ez nem teljesül, akkor is igaz, csak akkor az ábra nem áttekinthető.) Legyen Z=R-XWY.

Y

| | X | W | S | Y-S | Z |
|---|---|---|---|---|---|
| t | x | a | b | c | d |
| s | x | e | f | g | h |
| X→→Y miatt ∃u | x | a | f | g | d |

**W→S** miatt **b=f.**

Hasonlóan igazolható minden axióma helyessége. **q.e.d.**

# Többértékű függőségek

- **Állítás** (további levezetési szabályok):

  1. $X \rightarrow\rightarrow Y$ és $X \rightarrow\rightarrow V$ esetén $X \rightarrow\rightarrow YV$.
  2. $X \rightarrow\rightarrow Y$ és $WX \rightarrow\rightarrow V$ esetén $WX \rightarrow\rightarrow V-WY$.
  3. $X \rightarrow\rightarrow Y$ és $XY \rightarrow V$ esetén $X \rightarrow V-Y$.
  4. $X \rightarrow\rightarrow Y$ és $X \rightarrow\rightarrow V$ esetén $X \rightarrow\rightarrow Y \cap V$
     és $X \rightarrow\rightarrow V-Y$
     és $X \rightarrow\rightarrow Y-V$.

  **Bizonyítás:** A definíciókban szereplő sorok létezéséből, illetve értékek egyezőségéből az előző bizonyításhoz hasonlóan közvetlenül adódik. **q.e.d.**

# Többértékű függőségek

- **Állítás:** **$X \rightarrow\rightarrow Y$**-ből nem következik, hogy **$X \rightarrow\rightarrow A$,** ha **$A \in Y$.**
- **Bizonyítás:** A következő r tábla kielégíti az **$X \rightarrow\rightarrow AB$**-t, de nem elégíti ki az **$X \rightarrow\rightarrow A$**-t. **q.e.d.**

| X | A | B | C |
|---|---|---|---|
| x | a | b | c |
| x | e | f | g |
| x | a | b | g |
| x | e | f | c |

**$X \rightarrow\rightarrow A$** esetén ennek a sornak is benne kellene lenni a táblában.

| | | | |
|---|---|---|---|
| x | a | f | g |

# Többértékű függőségek

- **Állítás:** **X→→Y** és **Y→→V**-ből nem következik, hogy **X→→V.**
- **Bizonyítás:** A következő r tábla kielégíti az **X→→AB**-t, **AB→→BC**-t, de nem elégíti ki az **X→→BC**-t. **q.e.d.**

| X | A | B | C |
|---|---|---|---|
| x | a | b | c |
| x | e | f | g |
| x | a | b | g |
| x | e | f | c |

**X→→BC** esetén ennek a sornak is benne kellene lenni a táblában.

| | | | |
|---|---|---|---|
| x | e | b | c |

# Többértékű függőségek

- Funkcionális függőségek esetén az implikációs probléma hatékony megoldását az attribútum halmaz lezárására lehetett visszavezetni, mert

  $F \models X \rightarrow Y \Leftrightarrow F \vdash X \rightarrow Y \Leftrightarrow Y \subseteq X^*$,

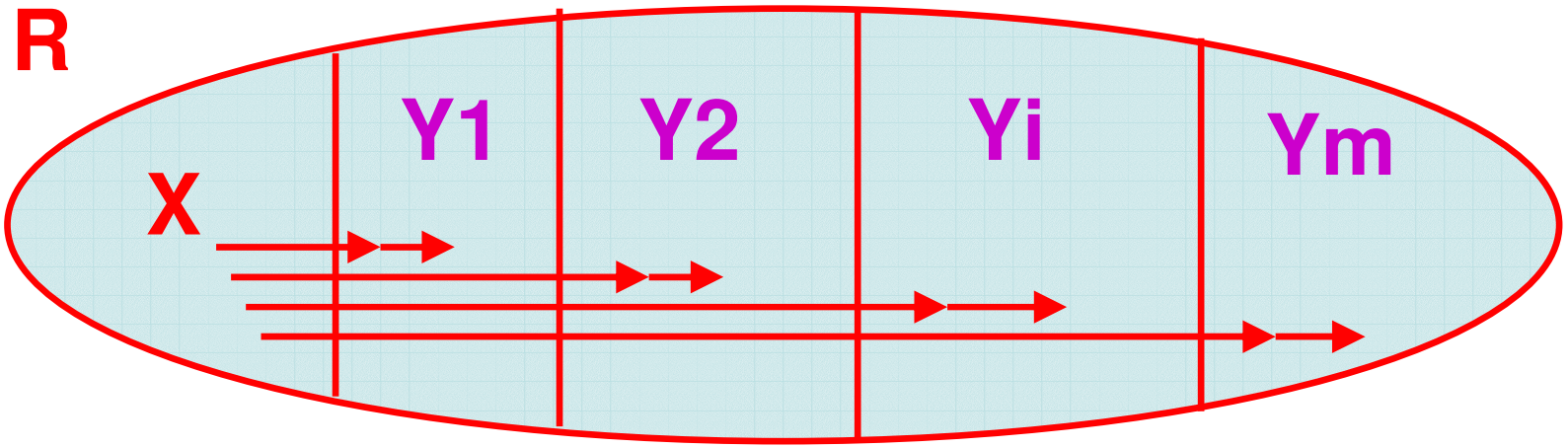
  ahol $X^* := \{A \mid F \vdash X \rightarrow A\}$
- A második ekvivalenciához fel kellett használni a dekompozíciós szabályt, de többértékű függőségek esetén nem érvényes a dekompozíciós szabály (a jobb oldalak nem vághatók szét). Így más módszer szükséges az összes levezethető függőség meghatározásához.
- Adott D esetén hogy lehet az összes $X \rightarrow\rightarrow Y$-t meghatározni, amely D következménye?
- Azt már tudjuk, hogy ha azonos X bal oldalú következményeket már találtunk, akkor a jobb oldalak egyesítésével, metszetével, különbségével újabbakra következtethetünk.
- **Ötlet:** Rögzített X bal oldal esetén keressük meg azokat a **bázishalmazokat**, amelyekből a fenti halmazműveletekkel az összes olyan jobb oldal összerakható, amire D következményét kapjuk.

# Többértékű függőségek

- **Definíció**: $X \subseteq R$ és adott D függőségi halmaz esetén **$B(X,D)=\{Y1,Y2,...,Ym\}$** az X függőségi bázisa D-re nézve,
  - ha B(X,D) az R-X diszjunkt partíciója,
  - és tetszőleges **$V \subseteq R-X$** esetén $D \models X \rightarrow\rightarrow V$ akkor és csak akkor, ha **V** előállítható **Yi** bázistagok egyesítéseként.
- Kérdések:
  - Létezik-e mindig függőségi bázis?
  - Egyértelmű-e?
  - Hatékony algoritmussal elő lehet-e állítani?
- **Következmény:** Egy tagú egyesítést véve, a függőségi bázis minden Yi tagjára teljesül, hogy $D \models X \rightarrow\rightarrow Yi$.

# Többértékű függőségek

- Ha **B(X,D)={Y1,Y2,...,Ym}** X függőségi bázisa:




- Tehát az összes **X→→V**, amire **V⊆R-X** és **D|= X→→V:**
  - **X→→Yi**
  - **X→→YiYj**
  - **X→→YiYjYk**
  - **....**
  - **X→→Y1Y2...Ym**

# Többértékű függőségek

- **Tétel:** $X \subseteq R$ és adott D függőségi halmaz esetén **$B(X,D)=\{Y1,Y2,...,Ym\}$** a halmazok sorrendjétől eltekintve egyértelmű, és meghatározható a következő naiv algoritmussal.

- $B(X,D)(0):=\{R-X\}$

  $B(X,D)(j)=\{W1,W2,...,Wk\}$ esetén vegyünk egy olyan V-t, amire $\boxed{D \models X \rightarrow\rightarrow V}$ és V nem áll elő a Wi-k egyesítéseként. (Ekkor V valódi módon belemetsz a Wi-k közül bizonyosakba.)

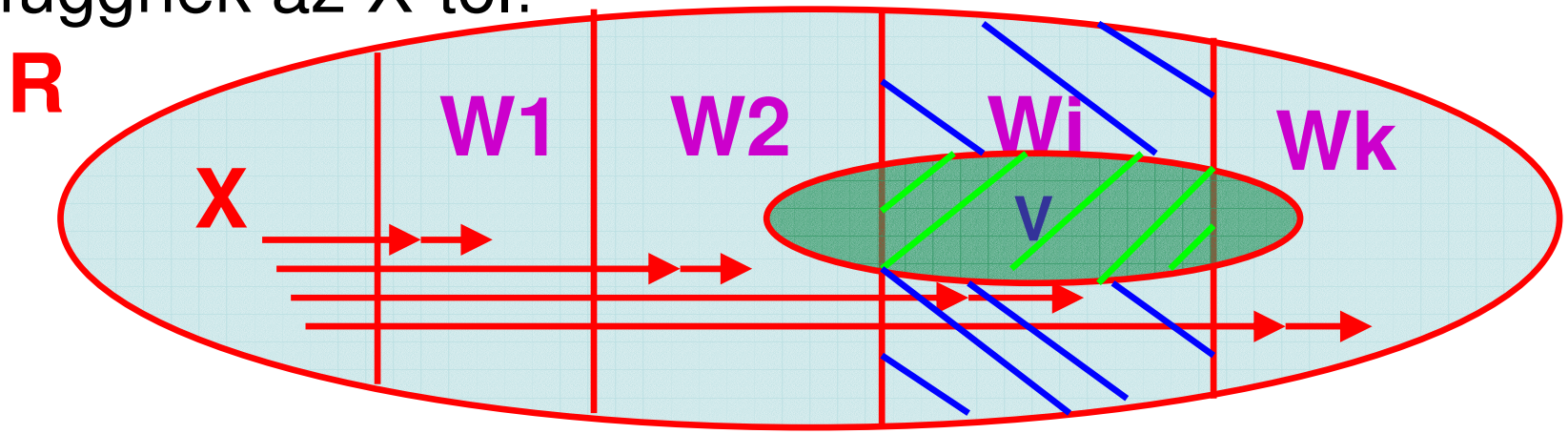
  $B(X,D)(j+1)$ –t úgy kapjuk, hogy **Wi**-ket lecseréljük **Wi-V** és **$Wi \cap V$** nem üres halmazokra.

  OUTPUT: $B(X,D)(n)$, ha $B(X,D)(n+1)=B(X,D)(n)$

**Bizonyítás:** Az algoritmus megáll, mert a partíciós tagokban csökken az attribútumok száma.

# Többértékű függőségek

- Elég belátni, hogy a partíciós tagok többértékűen függnek az X-től!




- **D|= X→→R-X**: mert **D|= X→∅**, A7 miatt **D|= X→→∅**, és A4 miatt **D|= X→→R-X∅** azaz **D|= X→→R-X**.
- **Indukció**: **D|= X→→Wi**: minden i-re.
- Mivel **D|= X→→V,** és a jobb oldalak metszete, különbsége is levezethető, így az új partíciós tagokra (**Wi-V** és **Wi∩V** halmazokra) is teljesül az implikáció:

  **D|= X→→ Wi-V** és **D|= X→→ Wi∩V. q.e.d.**

# Többértékű függőségek

- A naiv algoritmus nem hatékony, ráadásul ha tudnánk hogy kell olyan V-t találni, amelyre **D|= X→→V,** akkor nem is kellene függőségi bázis.
- Be fogjuk látni, hogy ha D-ben csak többértékű függőségek vannak, akkor már hatékonyan megtalálható a függőségi bázis.
- Hogy lehet megoldani egy vegyes függőségeket tartalmazó D-re vonatkozó implikációs problémát csak többértékű függőségekre vonatkozó függőségi bázissal?
- **D** vegyes függőségi halmaznak megfeleltetünk egy **M(D)** többértékű függőségeket tartalmazó halmazt.
- **M(D)**:= {**U→→V**| ha **U→→V** ∈ **D**}

  ∪ {**S→→Ai**| ha **S→A1...Ak** ∈ **D**}
- A következő tétel mutatja, hogy mi a kapcsolat **D** és **M(D)** között.

# Többértékű függőségek

- **Tétel:** Tegyük fel, hogy $X \cap Y = \emptyset$.
  $D \models X \rightarrow\rightarrow Y \Leftrightarrow Y$ előáll $B(X,M(D))$-hez tartozó partíciós tagok egyesítéseként.

- **Tétel:** Tegyük fel, hogy $A \notin X$.
  $D \models X \rightarrow A \Leftrightarrow$ **1.** $\{A\}$ halmaz szerepel $B(X,M(D))$-ben
  **2**. Van olyan $U \rightarrow V$ a $D$-ben, amelyre
  $A \notin U$ és $A \in V$.

- A két tételt nem bizonyítjuk, helyette megnézzük, hogy lehet $B(X,M(D))$ függőségi bázist meghatározni hatékonyan. Csak azt használjuk ki, hogy $M(D)$-ben csak többértékű függőségek szerepelnek.

# Többértékű függőségek

- **Tétel:** $X \subseteq R$ és adott **M**, csak többértékű függőségeket tartalmazó halmaz esetén **B(X,M) függőségi bázis meghatározható a következő algoritmussal.**

- **B(X,M)**(0):={R-X}

  **B(X,M)**(j)={W1,W2,...,Wk} esetén vegyünk egy olyan V-t, amire $\boxed{\mathbf{U} \rightarrow\rightarrow \mathbf{V} \in \mathbf{M},}$ és van olyan Wi, hogy $\boxed{\mathbf{U} \cap \mathbf{Wi} = \emptyset}$, és **V** valódi módon belemetsz a Wi-be.

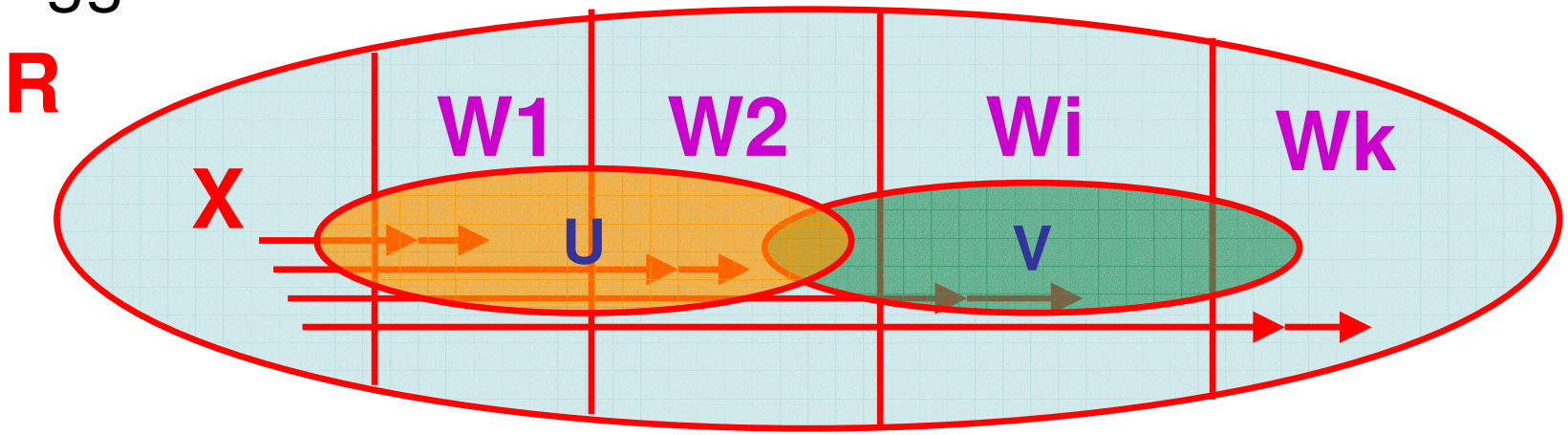
  **B(X,M)**(j+1) –t úgy kapjuk, hogy **Wi**-ket lecseréljük **Wi-V** és **Wi∩V** nem üres halmazokra.

  OUTPUT: **B(X,M)**(n), ha **B(X,M)**(n+1)=**B(X,M)**(n)

**Bizonyítás:** Az algoritmus megáll, mert a partíciós tagokban csökken az attribútumok száma.

# Többértékű függőségek

- Elég belátni, hogy a partíciós tagok többértékűen függnek az X-től!




- **M|= X→→R-X**: mert **M|= X→∅**, A7 miatt **M|= X→→∅**, és A4 miatt **M|= X→→R-X∅** azaz **M|= X→→R-X**.
- **Indukció**: **M|= X→→Wi**: minden i-re.
- A7 miatt **M|= X→→X,** és a jobb oldalak egyesíthetők, így **M|= X→→XW1...Wi-1Wi+1...Wk.**
- Mivel **U** része a jobb oldalnak, ezért A7 miatt **M|=XW1...Wi-1Wi+1...Wk→→U.**
- **U** és **V** választása miatt **M|= U→→V**

# Többértékű függőségek

- A tranzitív szabály, A7 miatt:

  **$M \models XW_1 \ldots W_{i-1} W_{i+1} \ldots W_k \rightarrow\rightarrow V-U$.**
- A tranzitív szabály, A7 miatt:

  **$M \models X \rightarrow\rightarrow (V-U)-XW_1 \ldots W_{i-1} W_{i+1} \ldots W_k$.**
  - A rajzon látszik, hogy a jobb oldal nem más, mint **$W_i \cap V$,** tehát kijött, hogy
  - **$M \models X \rightarrow\rightarrow W_i \cap V$.**
- Indukció miatt:
  - **$M \models X \rightarrow\rightarrow W_i$** és a jobb oldalak különbsége vehető, így:
  - **$M \models X \rightarrow\rightarrow W_i - (W_i \cap V)$,** azaz
  - **$M \models X \rightarrow\rightarrow W_i - V$.**
- Tehát az új partíciós tagokra is teljesül a függőség. **q.e.d.**
- Most azért volt bonyolultabb a bizonyítás, mert **X**-től különböző **U** szerepelt a bal oldalon.
- **Költség:** csak **M**-beli függőségeket kell venni, így **$O(|M|*|R|^3)$**

# Többértékű függőségek

- **Következmény:** $B(X,M(D))$ függőségi bázist meg lehet határozni hatékonyan, így a már idézett alábbi tétel miatt vegyes függőségekre is hatékonyan megoldható az implikációs probléma.
- **Tétel:** Tegyük fel, hogy $X \cap Y = \emptyset$.

  $D \models X \rightarrow\rightarrow Y \Leftrightarrow Y$ előáll $B(X,M(D))$-hez tartozó partíciós tagok egyesítéseként.
- **Tétel:** Tegyük fel, hogy $A \notin X$.

  $D \models X \rightarrow A \Leftrightarrow$ **1.** $\{A\}$ halmaz szerepel $B(X,M(D))$-ben

  **2.** Van olyan $U \rightarrow V$ a $D$-ben, amelyre

  $$A \notin U \text{ és } A \in V.$$
- **Megjegyzés:** Ha $X \cap Y = \emptyset$. nem teljesül, akkor is visszavezethető a fentiekre: Legyen $Y1 := Y - X$.

  $D \models X \rightarrow X \cap Y$ vagyis $D \models X \rightarrow\rightarrow X \cap Y$ mindig teljesül, azaz $D \models X \rightarrow\rightarrow Y \Leftrightarrow D \models X \rightarrow\rightarrow Y1$.

# Többértékű függőségek

- A **veszteségmentesség**, **függőségőrzés** definíciójában most **F** funkcionális függőségi halmaz helyett **D** függőségi halmaz többértékű függőségeket is tartalmazhat.
- Így például d=(R1,...,Rk) veszteségmentes dekompozíciója R-nek **D**-re nézve, akkor és csak akkor, ha minden $\mathbf{r \in SAT(D)}$ esetén $\mathbf{r = \Pi_{R1}(r) |\!\!><\!\!| \ldots |\!\!><\!\!| \Pi_{Rk}(r)}$

- A következő tétel miatt a veszteségmentesség implikációs problémára vezethető vissza, így hatékonyan eldönthető.

- **Tétel:** A d=(R1,R2) akkor és csak akkor **veszteségmentes** dekompozíciója R-nek, ha $\mathbf{D \models R1 \cap R2 \rightarrow\rightarrow R1 - R2}$.

# Többértékű függőségek

- **Bizonyítás.** Legyen **r∈SAT(D)**.

| R1-R2 | R1∩R2 | R2-R1 |
|---|---|---|
| a | b | e |
| f | b | c |
| a | b | c |

- A dekompozíció akkor és csak akkor **veszteségmentes**, ha a vetületek összekapcsolása része a relációnak, azaz **ab**$\in \Pi_{R1}$**(r)** és **bc**$\in \Pi_{R2}$**(r)** esetén **abc**∈**r,** ami pont azt jelenti, hogy a közös részen megegyező két sorhoz létezzen egy harmadik sor, ahogy a **R1∩R2→→R1-R2** többértékű függőség definíciójában szerepel. **q.e.d.**

# Többértékű függőségek

- A 4.normálforma definiálása előtt foglaljuk össze, hogy melyek a **triviális többértékű függőségek**, vagyis amelyek **minden relációban teljesülnek**.
- Mivel minden funkcionális függőség többértékű függőség is, így a triviális funkcionális egyben triviális többértékű függőség is.
- **$Y \subseteq X$** esetén **$X \rightarrow\rightarrow Y$ triviális többértékű függőség.**
- Speciálisan **$Y=\emptyset$** választással **$X \rightarrow\rightarrow \emptyset$** függőséget kapjuk, és alkalmazzuk a komplementer szabályt, azaz **$Z=R-X\emptyset$,** így az **$X \rightarrow\rightarrow R-X$** függőség is mindig teljesül, azaz:
- **$XY=R$** esetén **$X \rightarrow\rightarrow Y$ triviális többértékű függőség.**
- A szuperkulcs, kulcs definíciója változatlan, azaz **X szuperkulcsa R**-nek **D**-re nézve, ha **$D \models X \rightarrow R$.**
- A minimális szuperkulcsot **kulcsnak** hívjuk.

# Többértékű függőségek

- A 4.normálforma hasonlít a BCNF-re, azaz minden nem triviális többértékű függőség bal oldala szuperkulcs.
- **Definíció:** R **4NF**-ben van D-re nézve, ha **$XY \neq R$**, **$Y \not\subset X$,** és **$D \models X \rightarrow\rightarrow Y$** esetén **$D \models X \rightarrow R$.**
- **Definíció: d={R1,...,Rk}** dekompozíció **4NF**-ben van D-re nézve, ha minden Ri **4NF**-ben van $\Pi_{Ri}(D)$-re nézve.
- **Állítás:** Ha R **4NF**-ben van, akkor **BCNF**-ben is van.
- **Bizonyítás**. Vegyünk egy nem triviális **$D \models X \rightarrow A$** **funkcionális** függőséget. Ha **$XA=R$**, akkor **$D \models X \rightarrow R$**, ha **$XA \neq R$**, akkor a **$D \models X \rightarrow\rightarrow A$** nem triviális többértékű függőség és a **4NF** miatt **$D \models X \rightarrow R$.** **q.e.d.**
- **Következmény:** Nincs mindig **függőségőrző és veszteségmentes 4NF** dekompozíció.

# Többértékű függőségek

- **Veszteségmentes 4NF dekompozíció**t mindig tudunk készíteni a naiv BCNF dekomponáló algoritmushoz hasonlóan.
- **Naiv algoritmus veszteségmentes 4NF dekompozíció előállítására:**

  Ha R 4NF-ben van, akkor megállunk, egyébként van olyan nem triviális **$X \rightarrow\rightarrow Y$,** amely R-ben teljesül, de megsérti a 4NF-et, azaz X nem szuperkulcs. (Az is feltehető, hogy X és Y diszjunkt, mert különben Y-helyett az Y-X-et vehettük volna jobb oldalnak.) Ekkor **R helyett vegyük az (XY,R-Y)** dekompozíciót. **$XY \neq R$**, így *mindkét tagban csökken az attribútumok száma.*

  **$XY \cap (R-Y) = X \rightarrow\rightarrow Y = XY-(R-Y)$,** azaz a kéttagú dekompozícióknál bizonyított állítás miatt *veszteségmentes kettévágást kaptunk*. A kettévágásokat addig hajtjuk végre, amíg minden tag 4NF-ben nem lesz. Legrosszabb esetben a 2 oszlopos sémákig kell szétbontani, amelyek mindig 4NF-ben vannak, mivel nem lehet bennük nem triviális többértékű függőség.