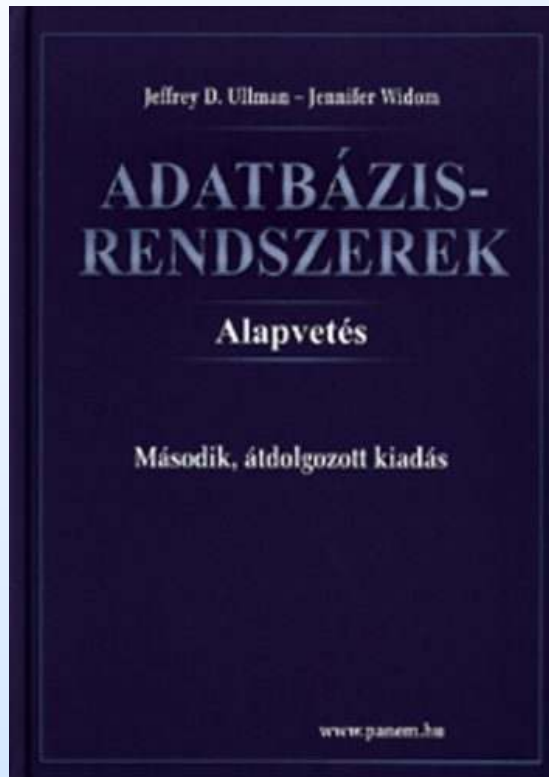


# Relációs adatbázisok tervezése

## 4.rész (többértékű függőségek)



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

3.6. Többértékű függőségek,  
- Negyedik normálforma  
- Funkcionális és többértékű függőségek következtetése

(Jeffrey D. Ullman 2007 előadásdiái alapján, Benczúr András, Kiss Attila és Kósa Balázs előadásainak felhasználásával, Hajas Csilla)

# Többértékű függőségek és 4NF

- ◆ Hasonló utat járunk be, mint a funkcionális függőségek esetén:
  - ▶ Definiáljuk a többértékű függőséget
  - ▶ implikációs probléma
  - ▶ axiomatizálás
  - ▶ levezethető függőségek hatékony meghatározása (lezárás helyett a séma partíciója függőségi bázisa)
  - ▶ veszteségmentes dekompozíció
  - ▶ 4. normálforma
  - ▶ veszteségmentes 4NF dekompozíció előállítás

# A TÉF definíciója

- ◆ A *többértékű függőség* (TÉF): az  $R$  reláció fölött  $X \twoheadrightarrow Y$  teljesül: ha bármely két sorra, amelyek megegyeznek az  $X$  minden attribútumán, az  $Y$  attribútumaihoz tartozó értékek felcserélhetőek, azaz a keletkező két új sor  $R$ -beli lesz.
- ◆ Más szavakkal:  $X$  minden értéke esetén az  $Y$ -hoz tartozó értékek függetlenek az  $R$ - $X$ - $Y$  értékeitől.

# Példa: TÉF

Sörivók(név, cím, tel, kedveltSörök)

- ◆ A sörivók telefonszámai függetlenek az általuk kedvelt söröktől.
  - ▶ név->->tel és név ->->kedveltSörök.
- ◆ Így egy-egy sörivó minden telefonszáma minden általa kedvelt sörrel kombinációban áll.
- ◆ Ez a jelenség független a funkcionális függőségektől.
  - ▶ itt a név->cím az egyetlen FF.

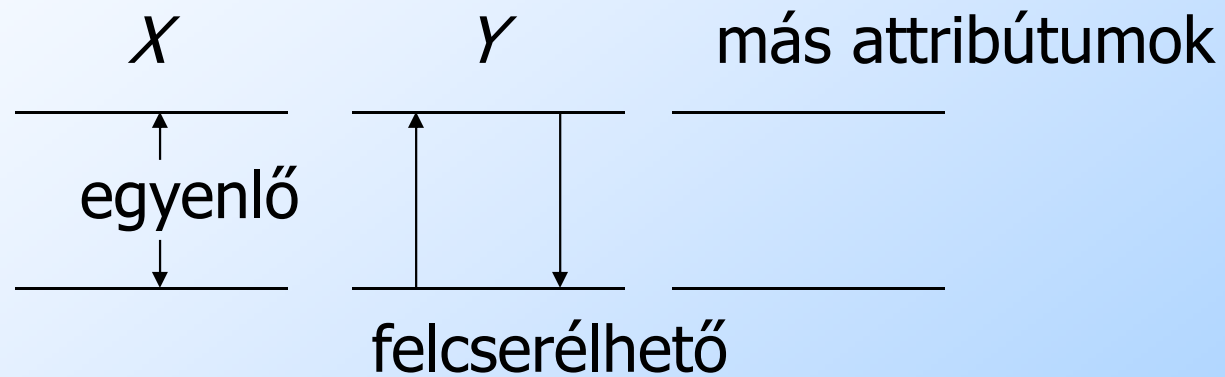
# A név->->tel által implikált sorok

Ha ezek a soraink vannak:

név	cím	tel	kedveltSörök
sue	a	p1	b1
sue	a	p2	b2
sue	a	p2	b1
sue	a	p1	b2

Akkor ezeknek a soroknak is szerepelnie kell.

# Az $X \rightarrow Y$ TÉF képe



# Többértékű függőségek

- ◆ **Definíció:**  $X, Y \subseteq R$ ,  $Z := R - XY$  esetén  $X \twoheadrightarrow Y$  **többértékű függőség. (tf)**
- ◆ A függőség akkor teljesül egy táblában, ha bizonyos mintájú sorok létezése garantálja más sorok létezését.
- ◆ A formális definíciót az alábbi ábra szemlélteti.
- ◆ Ha létezik **t** és **s** sor, akkor **u** és **v** soroknak is létezniük kell, ahol az azonos szimbólumok azonos értékeket jelölnek.

	<b>X</b>	<b>Y</b>	<b>Z</b>
t	x	y1	z1
s	x	y2	z2
$\exists u$	x	y1	z2
$\exists v$	x	y2	z1

# Többértékű függőségek

**Definíció (Formálisan):** Egy R sémájú  $r$  reláció kielégíti az  $X \twoheadrightarrow Y$  függőséget, ha  $t, s \in r$  és  $t[X] = s[X]$  esetén létezik olyan  $u, v \in r$ , amelyre  $u[X] = v[X] = t[X] = s[X]$ ,  $u[Y] = t[Y]$ ,  $u[Z] = s[Z]$ ,  $v[Y] = s[Y]$ ,  $v[Z] = t[Z]$ .

**Állítás:** Elég az  $u, v$  közül csak az egyik létezését megkövetelni.

	X	Y	Z
t	x	y1	z1
s	x	y2	z2
$\exists u$	x	y1	z2



# TÉF szabályok

- ◆ Minden FF TÉF.

- ▶ Ha  $X \rightarrow Y$  és két sor megegyezik  $X$ -en,  $Y$ -on is megegyezik, emiatt ha ezeket felcseréljük, az eredeti sorokat kapjuk vissza, azaz:  $X \rightarrow \rightarrow Y$ .

- ◆ *Komplementálás* : Ha  $X \rightarrow \rightarrow Y$  és  $Z$  jelöli az összes többi attribútum halmazát, akkor  $X \rightarrow \rightarrow Z$ .

# Nem tudunk darabolni

- ◆ Ugyanúgy, mint az FF-ek esetében, a baloldalakat nem „bánthatjuk” általában.
- ◆ Az FF-ek esetében a jobboldalakt felbonthattuk, míg ebben az esetben ez sem tehető meg.

# Példa: többattribútumos jobboldal

Sörivők(név, tTársaság, tel, kedveltSörök, gyártó)

- ◆ Egy sörivónak több telefonja lehet, minden számot két részre osztunk: tTársaság (pl. Vodafone) és a maradék hét számjegy.
- ◆ Egy sörivó több sört is kedvelhet, mindegyikhez egy-egy gyártó tartozik.

## Példa folytatás

- ◆ Mivel a tTársaság-tel kombinációk függetlenek a kedveltSörök-gyártó kombinációtól, azt várjuk, hogy a következő FÉK-ek teljesülnek:

név ->-> tTársaság tel

név ->-> kedveltSörök gyártó

# Példa adat

Egy lehetséges előfordulás, ami teljesíti az iménti FÉK-et:

név	tTársaság	tel	kedveltS	gyártó
Sue	30	555-1111	Bud	A.B.
Sue	20	555-1111	WickedAle	Pete's
Sue	70	555-9999	Bud	A.B.
Sue	70	555-9999	WickedAle	Pete's

Ugyanakkor sem a  $\text{név} \rightarrow \text{tTársaság}$  sem a  $\text{név} \rightarrow \text{tel}$  függőségek nem teljesülnek.

# Többértékű függőségek

## ◆ Axiomatizálás

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

# Többértékű függőségek

## ◆ Jelölés a továbbiakban:

- ▶ **F** funkcionális függőségek halmaza
- ▶ **M** többértékű függőségek halmaza
- ▶ **D** vegyes függőségek (funkcionális és többértékű függőségek) halmaza

## ◆ **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.**

## Többértékű függőségek

- ◆ **Állítás:**  $X \twoheadrightarrow Y$ -ből **nem következik**, hogy  $X \twoheadrightarrow A$ , ha  $A \in Y$ . (A jobb oldalak nem szedhetők szét!)
- ◆ **Bizonyítás:** A következő r tábla kielégíti az  $X \twoheadrightarrow AB$ -t, de nem elégíti ki az  $X \twoheadrightarrow A$ -t. q.e.d.

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

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

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



## Többértékű függőségek

- ◆ **Állítás:**  $X \twoheadrightarrow Y$  és  $Y \twoheadrightarrow V$ -ből **nem következik**, hogy  $X \twoheadrightarrow V$ . (A szokásos tranzitivitás nem igaz általában!)
- ◆ **Bizonyítás:** A következő r tábla kielégíti az  $X \twoheadrightarrow AB$ -t,  $AB \twoheadrightarrow BC$ -t, de nem elégíti ki az  $X \twoheadrightarrow BC$ -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	e	b	c
---	---	---	---

$X \twoheadrightarrow BC$  esetén ennek a sornak is benne kellene lenni a táblában.

# Többértékű függőségek

- ◆ A **veszteségmentesség, függőségörzés** definíciójában most **F** funkcionális függőségi halmaz helyett **D** függőségi halmaz többértékű függőségeket is tartalmazhat.
- ◆ Így például  **$d=(R_1, \dots, R_k)$**  veszteségmentes dekompozíciója **R**-nek **D**-re nézve, akkor és csak akkor, ha minden **D**-t kielégítő **r** tábla esetén  
 **$r = \Pi_{R_1}(r) \mid \dots \mid \Pi_{R_k}(r)$**
- ◆ A következő tétel miatt a **veszteségmentesség implikációs problémára vezethető vissza**, így hatékonyan eldönthető.
- ◆ **Tétel:** A  **$d=(R_1, R_2)$**  akkor és csak akkor **veszteségmentes** dekompozíciója **R**-nek, ha  **$D \mid \text{--- } R_1 \cap R_2 \rightarrow \rightarrow R_1 - R_2$** .

## 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.
  1.  $Y \subseteq X$  esetén  $X \twoheadrightarrow Y$  **triviális többértékű függőség**.
- ◆ Speciálisan  $Y = \emptyset$  választással  $X \twoheadrightarrow \emptyset$  függőséget kapjuk, és alkalmazzuk a komplementer szabályt, azaz  $Z = R - X\emptyset$ , így az  $X \twoheadrightarrow R - X$  függőség is mindig teljesül, azaz:
  2.  $XY = R$  esetén  $X \twoheadrightarrow Y$  **triviális többértékű függőség**.
- ◆ A superkulcs, kulcs definíciója változatlan, azaz  $X$  **superkulcsa**  $R$ -nek  $D$ -re nézve, ha  $D \mid\!\!\!\mid X \rightarrow R$ .
- ◆ A minimális superkulcsot **kulcsnak** hívjuk.

## Többértékű függőségek

- ◆ A 4.normálforma hasonlít a BCNF-re, azaz minden nem triviális többértékű függőség bal oldala superkulcs.
- ◆ **Definíció:** R **4NF**-ben van D-re nézve, ha  $XY \neq R$ ,  $Y \not\subseteq X$ , és  $D \models X \twoheadrightarrow Y$  esetén  $D \models X \rightarrow R$ .
- ◆ **Definíció:**  $d = \{R_1, \dots, R_k\}$  dekompozíció **4NF**-ben van D-re nézve, ha minden  $R_i$  **4NF**-ben van  $\Pi_{R_i}(D)$ -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  $D \models X \twoheadrightarrow A$  nem triviális többértékű függőség és a **4NF** miatt  $D \models X \rightarrow R$ . q.e.d.
- ◆ **Következmény:** Nincs mindig **függőségörző** és **veszteségmentes 4NF** dekompozíció.

## Többértékű függőségek

- ◆ **Veszteségmentes 4NF** dekompozíciót 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  $X \twoheadrightarrow Y$ , amely R-ben teljesül, de **megsérti a 4NF-et**, azaz X nem superkulcs.  
Ekkor **R helyett vegyük az (XY, R-Y)** dekompozíciót.  
A kettévágásokat addig hajtjuk végre, amíg minden tag 4NF-ben nem lesz. **ALGORITMUS VÉGE.**
- ◆ Az is feltehető, hogy X és Y diszjunkt, mert különben Y helyett az Y-X-et vehettük volna jobb oldalnak.
- ◆  $XY \neq R$ , így *mindkét tagban csökken az attribútumok száma.*
- ◆  $XY \cap (R-Y) = X \twoheadrightarrow Y = XY - (R-Y)$ , azaz a kéttagú dekompozícióknál bizonyított állítás miatt *veszteségmentes kettévágást kaptunk.*
- ◆ Legrosszabb esetben a 2 oszlopos sémáig kell szétbontani, amelyek mindig 4NF-ben vannak, mivel nem lehet bennük nem triviális többértékű függőség.

# Negyedik normálforma

- ◆ A TÉF-ek okozta redundanciát a BCNF nem szünteti meg.
- ◆ A megoldás: a negyedik normálforma!
- ◆ A negyedik normálformában (4NF), amikor dekomponálunk, a TÉF-eket úgy kezeljük, mint az FF-eket, a kulcsok megtalálásánál azonban nem számítanak.

# 4NF definíció

- ◆ Egy  $R$  reláció **4NF**-ben van ha:  
minden  $X \twoheadrightarrow Y$  nemtriviális FÉK  
esetén  $X$  superkulcs.
- ◆ **Nemtriviális TEF** :
  1.  $Y$  nem részhalmaza  $X$ -nek,
  2.  $X$  és  $Y$  együtt nem adják ki az összes attribútumot.
- ◆ A superkulcs definíciója ugyanaz marad,  
azaz csak az FF-ektől függ.

# BCNF versus 4NF

- ◆ Kiderült, hogy minden  $X \twoheadrightarrow Y$  FF  $X \twoheadrightarrow - \twoheadrightarrow Y$  TÉF is.
- ◆ Így, ha  $R$  4NF-ben van, akkor BCNF-ben is.
  - ▶ Mert minden olyan FF, ami megsérti a BCNF-t, a 4NF-t is megsérti.
- ◆ De  $R$  lehet úgy BCNF-ben, hogy közben nincs 4NF-ben.



# Dekompozíció és 4NF

- ◆ H  $X \twoheadrightarrow Y$  megsérti a 4NF-t, akkor  $R$ -t ugyanúgy dekomponáljuk, mint a BCNF esetén.
  1.  $XY$  az egyik dekomponált reláció.
  2. Az  $Y - X$ -be nem tartozó attribútumok a másik.

# Példa: 4NF dekompozíció

Sörivók(név, cím, tel, kedveltSörök)

FF: név -> cím

FÉK-ek: név ->-> tel

név ->-> kedveltSörök

- ◆ Kulcs {név, tel, kedveltSörök}.
- ◆ Az összes függőség megsérti 4NF-et.

# Példa folytatás

- ◆ Dekompozíció **név** -> **cím** szerint:

## 1. Sörivók1(név, cím)

- ◆ Ez 4NF-beli; az egyetlen függőség **név**-> **cím**.

## 2. Sörivók2(név, tel, kedveltSörök)

- ◆ Nincs 4NF-ben. A **név** ->-> **tel** és **név** ->-> **kedveltSörök** függőségek teljesülnek. A három attribútum együtt kulcs (mivel nincs nemtriviális FF).

# Példa: Sörivók2 dekompozíciója

- ◆ Mind a  $\text{név} \rightarrow \rightarrow \text{tel}$ , mind a  $\text{név} \rightarrow \rightarrow \text{kedveltSörök}$  szerinti dekompozíció ugyanazt eredményezi:
  - ▶  $\text{Sörivók3}(\underline{\text{név}}, \underline{\text{tel}})$
  - ▶  $\text{Sörivók4}(\underline{\text{név}}, \underline{\text{kedveltSörök}})$