

# Adatbázisok 1.

## Relációs adatbázis tervezés – 5. rész

Funkcionális függőségek

Felbontások

Normálformák

# A harmadik normálforma (*third normal form*) -- motiváció

- Bizonyos FF halmazok esetén a felbontáskor elveszíthetünk függőségeket.
- $AB \rightarrow C$  és  $C \rightarrow B$ .
  - Példa:  $A = \text{f\_cím}$ ,  $B = \text{város}$ ,  $C = \text{mozi}$ .
- Két kulcs van:  $\{A, B\}$  és  $\{A, C\}$ .
- $\{\text{f\_cím}, \text{város}\}, \{\text{f\_cím}, \text{mozi}\}$
- $C \rightarrow B$  megsérti a BCNF-t, tehát  $AC$ ,  $BC$ -re dekomponálunk. [ $\text{mozi} \rightarrow \text{város}$ , nem superkulcs  $C$ ]

# FF-ek kikényszerítése

- A probléma az, hogy  $AC$  és  $BC$  sémákkal nem tudjuk kikényszeríteni  $AB \rightarrow C$  függőséget.
- Példa  $A = f\_cím$ ,  $B = város$ ,  $C = mozi$ , a következő dián.

# Egy kikényszeríthetetlen (*unenforceable*) FF

F_cím	mozi
Antz	Guild
Antz	Park

város	mozi
Cambridge	Guild
Cambridge	Park

Kapcsoljuk össze a sorokat (mozi).

F_cím	város	mozi
Antz	Cambridge	Guild
Antz	Cambridge	Park

A szétbontott relációkban egyik FF sem sérül, az eredményben az **F\_cím város -> mozi** nem teljesül.

# A probléma megoldása: 3NF

- 3. normálformában (3NF) úgy módosul a BCNF feltétel, hogy az előbbi esetben nem kell dekomponálnunk.
- Egy attribútum **prím, azaz elsődleges attribútum** (*prime attribute*), ha legalább egy kulcsnak eleme.
- $X \rightarrow A$  nem-trivi. FF, megsérti 3NF-t akkor és csak akkor, ha  $X$  nem superkulcs és  $A$  nem prím.
- minden nem triviális függőségre igaz, hogy bal oldala superkulcs, vagy jobb oldala csak elsődleges attribútumokat tartalmaz
- 3NF feltétel és a BCNF feltétel közötti különbség a „vagy jobb oldala csak elsődleges attribútumokat tartalmaz”

## Példa: 3NF

- A problematikus esetben az  $AB \rightarrow C$  és  $C \rightarrow B$  FF-ek esetén a kulcsok  $AB$  és  $AC$ .
- Ezért  $A$ ,  $B$  és  $C$  mindegyike prím.
- Habár  $C \rightarrow B$  megsérti a BCNF feltételét, 3NF feltételét már nem sérti meg.

# Miért hasznos 3NF és BCNF?

- A dekompozícióknak három fontos tulajdonsága lehet:
  1. *Veszteségmentes összekapcsolás (lossless join)*, információ visszaállíthatóság
  2. *Függőségek megőrzése (dependency preservation)*
  3. *Anomáliák kiküszöbölése (elimination of anomalies)*

# 3NF és BCNF -- folytatás

- Az (1) tulajdonság teljesül a BCNF esetében.
- A 3NF (1) és (2)-t is teljesíti, de maradhat benne anomália (általában ez nem akkora baj)
- A BCNF esetén (2) sérülhet, viszont (FF-ek okozta) anomália nem lehet benne
  - Az **F\_cím - város - mozi** erre volt egy példa.

F_cím	város	mozi
Antz	Cambridge	Guild
A Bug's Life	Cambridge	Guild



# Minimális bázis létrehozása

- Minimális bázis (*minimal basis*)
  1. Jobb oldalak szétvágása.
  2. Próbáljuk törölni az FF-eket egymás után. Ha a megmaradó FF-halmaz nem ekvivalens az eredetivel, akkor nem törölhető az épp aktuális FF.
  3. Egymás után próbáljuk csökkenteni a baloldalakat, és megnézzük, hogy az eredetivel ekvivalens FF-halmazt kapunk-e.

## 3NF-re bontás – (2)

- A minimális bázis minden FF-re megad egy sémát a felbontásban.
  - A séma a jobb- és baloldalak uniója lesz ( $X \rightarrow A$  FF,  $XA$  séma) .
- Ha a minimális bázis FF-jei által meghatározott sémák ( $X \rightarrow A$  FF,  $XA$  séma) között nincs *superkulcs*, akkor hozzáadunk a felbontáshoz egy olyan *sémát*, amely maga egy *kulcs* az R relációra.

## Példa: 3NF felbontás

- A reláció:  $R = ABCD$ .
- FF-ek:  $A \rightarrow B$  és  $A \rightarrow C$ .
- **Felbontás**:  $AB$  és  $AC$  az FF-ekből és  $AD$ -t is hozzá kell venni, mert  $AB$ ,  $AC$  egyike sem kulcs.

# Miért működik?

- **Megőrzi a függőségeket:** minden FF megmarad a minimális bázisból.
- **Veszteségmentes összekapcsolás:** a CHASE algoritmussal ellenőrizhető (a kulcsból létrehozott séma itt lesz fontos).
- **3NF:** a minimális bázis tulajdonságaiból következik.

# Minimális bázist kiszámító algoritmus

Jelölje  $F^+$  az  $F$  függőségi halmazból következő függőségek halmazát.

1. Kezdetben  $G$  az üreshalmaz.
2. Minden  $X \rightarrow Y \in F$  helyett vegyük az  $X \rightarrow A$  függőségeket, ahol  $A \in Y - X$ ).

**Megjegyzés:** Ekkor minden  $G$ -beli függőség  $X \rightarrow A$  alakú, ahol  $A$  attribútum.

3. Minden  $X \rightarrow A \in G$ -re, ha  $X \rightarrow A \in (G - \{X \rightarrow A\})^+$ , vagyis ha elhagyjuk az  $X \rightarrow A$  függőséget  $G$ -ből, az még mindig következik a maradékból, akkor  $G := G - \{X \rightarrow A\}$ .

**Megjegyzés:** Végül nem marad több elhagyható függőség.

4. Minden  $X \rightarrow A \in G$ -re, amíg van olyan  $B \in X$ -re, hogy  $A \in (X - B)^+$  a  $G$ -szerint, vagyis  $(X - B) \rightarrow A$  teljesül, akkor  $X := X - B$ .

**Megjegyzés:** E lépés után minden baloldal minimális lesz.