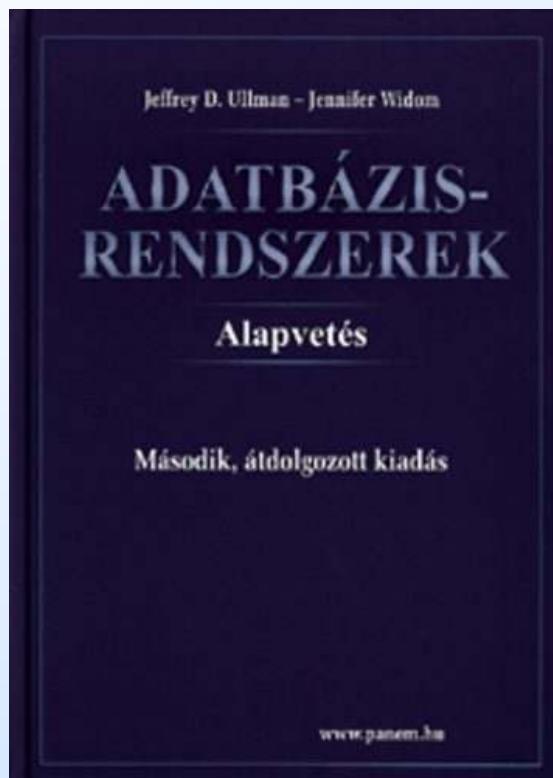


Relációs adatbázisok tervezése

2.rész (Boyce-Codd normálforma)



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

- ### 3.3. Relációs adatbázissémák tervezése
- Anomáliák. Relációk felbontása.
 - Boyce-Codd normálforma

(Jeffrey D. Ullman, 2007 slides alapján
Dr. Kiss Attila előadásainak felhasználásával)

Relational Schema Design

- ◆ Goal of relational schema design is to avoid anomalies and redundancy.
 - ▷ *Update anomaly* : one occurrence of a fact is changed, but not all occurrences.
 - ▷ *Deletion anomaly* : valid fact is lost when a tuple is deleted.

Example of Bad Design

Drinkers(name, addr, beersLiked, manf, favBeer)

name	addr	beersLiked	manf	favBeer
Janeway	Voyager	Bud	A.B.	WickedAle
Janeway	???	WickedAle	Pete's	???
Spock	Enterprise	Bud	???	Bud

Data is redundant, because each of the ???'s can be figured out by using the FD's name -> addr favBeer and beersLiked -> manf.

This Bad Design Also Exhibits Anomalies

name	addr	beersLiked	manf	favBeer
Janeway	Voyager	Bud	A.B.	WickedAle
Janeway	Voyager	WickedAle	Pete's	WickedAle
Spock	Enterprise	Bud	A.B.	Bud

- **Update anomaly:** if Janeway is transferred to *Intrepid*, will we remember to change each of her tuples?
- **Deletion anomaly:** If nobody likes Bud, we lose track of the fact that Anheuser-Busch manufactures Bud.

Boyce-Codd Normal Form

- ◆ We say a relation R is in ***BCNF*** if whenever $X \rightarrow Y$ is a nontrivial FD that holds in R , X is a superkey.
 - ▷ Remember: *nontrivial* means Y is not contained in X .
 - ▷ Remember, a *superkey* is any superset of a key (not necessarily a proper superset).

Example

Drinkers(name, addr, beersLiked, manf, favBeer)

FD's: name->addr favBeer, beersLiked->manf

- ◆ Only key is {name, beersLiked}.
- ◆ In each FD, the left side is *not* a superkey.
- ◆ Any one of these FD's shows *Drinkers* is not in BCNF

Another Example

Beers(name, manf, manfAddr)

FD's: name->manf, manf->manfAddr

- ◆ Only key is {name} .
- ◆ name->manf does not violate BCNF, but manf->manfAddr does.

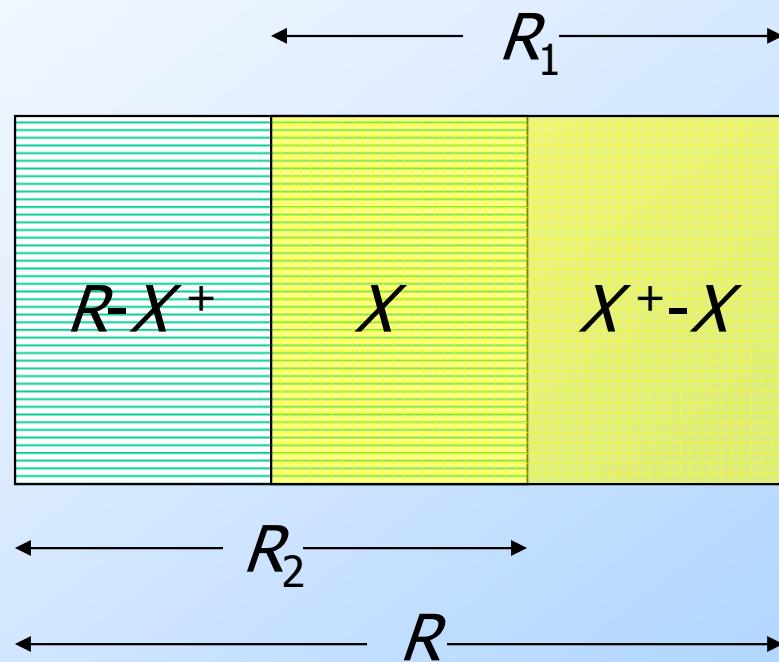
Decomposition into BCNF

- ◆ Given: relation R with FD's F .
- ◆ Look among the given FD's for a BCNF violation $X \rightarrow Y$.
 - If any FD following from F violates BCNF, then there will surely be an FD in F itself that violates BCNF.
- ◆ Compute X^+ .
 - Not all attributes, or else X is a superkey.

Decompose R Using $X \rightarrow Y$

- ◆ Replace R by relations with schemas:
 1. $R_1 = X^+$.
 2. $R_2 = R - (X^+ - X)$.
- ◆ *Project* given FD's F onto the two new relations.

Decomposition Picture



Example: BCNF Decomposition

Drinkers(name, addr, beersLiked, manf, favBeer)

$F = \text{name} \rightarrow \text{addr}$, $\text{name} \rightarrow \text{favBeer}$,
 $\text{beersLiked} \rightarrow \text{manf}$

- ◆ Pick BCNF violation $\text{name} \rightarrow \text{addr}$.
- ◆ Close the left side:
 $\{\text{name}\}^+ = \{\text{name}, \text{addr}, \text{favBeer}\}$.
- ◆ Decomposed relations:
 1. Drinkers1(name, addr, favBeer)
 2. Drinkers2(name, beersLiked, manf)

Example -- Continued

- ◆ We are not done; we need to check Drinkers1 and Drinkers2 for BCNF.
- ◆ Projecting FD's is easy here.
- ◆ For $\text{Drinkers1}(\underline{\text{name}}, \text{addr}, \text{favBeer})$, relevant FD's are $\text{name} \rightarrow \text{addr}$ and $\text{name} \rightarrow \text{favBeer}$.
 - ▶ Thus, $\{\text{name}\}$ is the only key and Drinkers1 is in BCNF.

Example -- Continued

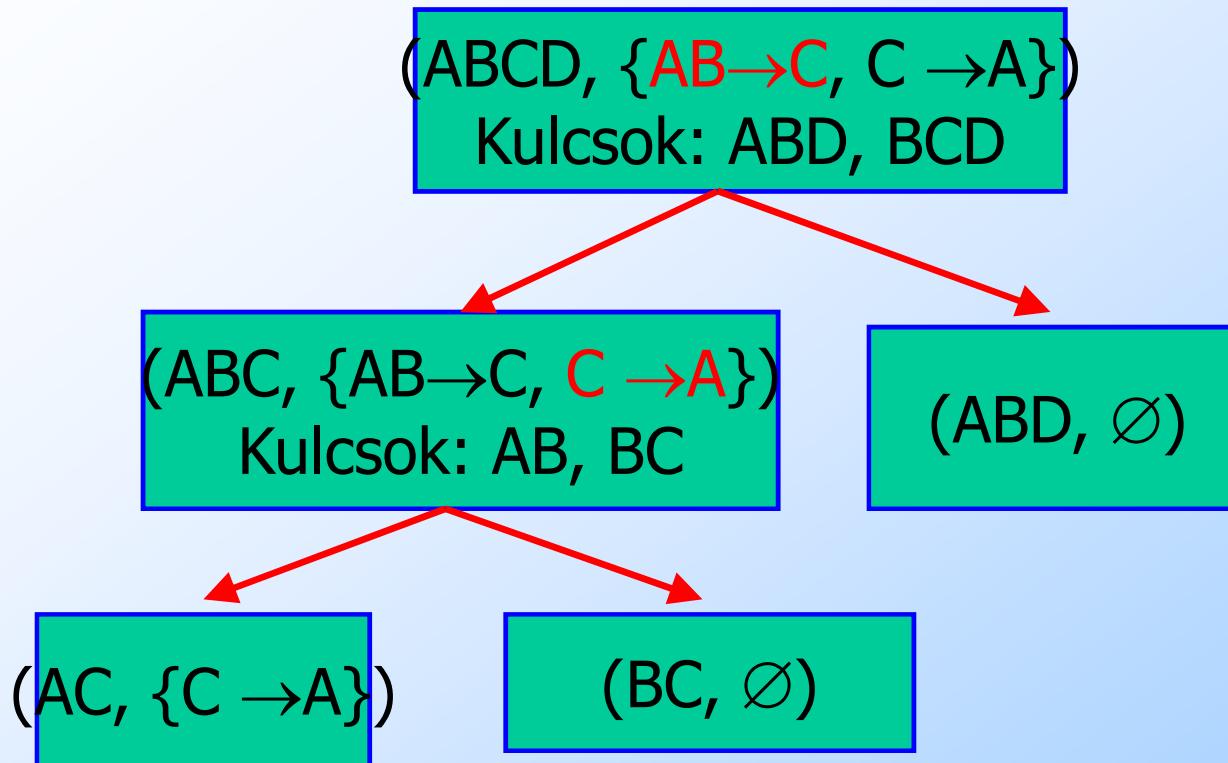
- ◆ For $\text{Drinkers2}(\underline{\text{name}}, \underline{\text{beersLiked}}, \text{manf})$, the only FD is $\text{beersLiked} \rightarrow \text{manf}$, and the only key is $\{\text{name}, \text{beersLiked}\}$.
 - ▶ Violation of BCNF.
- ◆ $\text{beersLiked}^+ = \{\text{beersLiked}, \text{manf}\}$, so we decompose Drinkers2 into:
 1. $\text{Drinkers3}(\underline{\text{beersLiked}}, \text{manf})$
 2. $\text{Drinkers4}(\underline{\text{name}}, \underline{\text{beersLiked}})$

Example -- Concluded

- ◆ The resulting decomposition of *Drinkers* :
 1. *Drinkers1*(name, addr, favBeer)
 2. *Drinkers3*(beersLiked, manf)
 3. *Drinkers4*(name, beersLiked)
- ◆ Notice: *Drinkers1* tells us about drinkers, *Drinkers3* tells us about beers, and *Drinkers4* tells us the relationship between drinkers and the beers they like.

BCNF-ra való felbontás

$R=ABCD$, $F=\{AB \rightarrow C, C \rightarrow A\}$



Tehát $d=(AC, BC, ABD)$ veszteségmentes BCNF dekompozíció.
(\emptyset azt jelenti, hogy csak a triviális függőségek teljesülnek a sémában.)

Tankönyv 3.5.2. feladata (111.o.)

- ◆ **Órarend adatbázis:** Kurzus(**K**), Oktató(**O**), Időpont(**I**), Terem(**T**), Diák(**D**), Csoport(**C**)

- ◆ **Feltételek:**

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

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

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

Egy időben egy diák csak egy helyen lehet: $ID \rightarrow T$.

Egy diák egy kurzis egy csoportjába jár: $KD \rightarrow C$.

- ◆ **R=KOITDJ F=** $\{K \rightarrow O, IT \rightarrow K, IO \rightarrow T, ID \rightarrow T, KD \rightarrow C\}$

- ◆ **Feladat:** Állítsuk elő (**R**, **F**) BCNF-ra való felbontását!