

# NORMALIZÁLÁS

Funkcionális függés

Redundancia

1NF, 2NF, 3NF

# FUNKCIONÁLIS FÜGGŐSÉG

- ▶ Legyen adott  $R(A_1, \dots, A_n)$  relációséma, valamint  $P, Q \subseteq \{A_1, \dots, A_n\}$  (magyarán  $P$  és  $Q$  a séma attribútumainak részhalmazai)
- ▶  $P$ -től funkcionálisan függ  $Q$ , jelölésben  $P \rightarrow Q$ , ha bármely  $R$  feletti  $T$  táblában valahányszor két sor megegyezik  $P$ -n, akkor megegyezik  $Q$ -n is.
- ▶ A  $P \rightarrow Q$  függés **triviális**, ha  $Q \subseteq P$ , vagy **teljesen nemtriviális**, ha  $P$  és  $Q$  metszete üres.

EHA	Név	Lakcím	Tárgy	Jegy
MINTAAT.SZE	Minta Áron	Szeged, Egy u. 2.	Adatbázisok	4
KELPEET.SZE	Kelep Elek	Sándorfalva, Fő tér 9.	Logika	3
MINTAAT.SZE	Minta Áron	Szeged, Egy u. 2.	Logika	5

A séma tulajdonsága!

# FUNKCIONÁLIS FÜGGÉS FOLYT. REDUNDANCIA

- ▶ Például egy függés: {EHA} → {Név, Lakcím}

EHA	Név	Lakcím	Tárgy	Jegy
MINTAAT.SZE	Minta Áron	Szeged, Egy u. 2.	Adatbázisok	4
KELPEET.SZE	Kelep Elek	Sándorfalva, Fő tér 9.	Logika	3
MINTAAT.SZE	Minta Áron	Szeged, Egy u. 2.	Logika	5

- ▶ **Redundanciáról** akkor beszélünk, ha valamely adatot feleslegesen többszörösen tárolunk el. Például, itt egy illető lakcíme feleslegesen többször szerepel. A fő gond ezzel, hogy adatok aktualizálásakor (pl. lakcímváltozás) minden helyen át kellene írni a lakcímet.
- ▶ A redundancia felszámolását a séma dekompozíciójával, azaz több sémára bontásával érjük el

# FELADAT

5.1 Mik a függőségek az alábbi adatbázissémában?

TANÁR (T.Azonosító, Név)

DIÁK (D.Azonosító, Név, Lakcím, SZ.Azonosító)

SZAK (SZ.Azonosító, Kar, Szak név)

TANÍTJA (T.Azonosító, D.Azonosító, Tantárgy)

# FELADAT

5.2 Tekintsük a korábban látott sémát:

Hallgató (EHA, Név, Lakcím, Tárgy, Jegy)

Végezzük el a tábla dekompozícióját úgy, hogy a kapott eredményben ne legyen már redundancia.

Hallgató (EHA, Név, Lakcím)

Eredmény (EHA, Tárgy, Jegy)

# CÉLKITŰZÉS

- ▶ Az adatok

- ▶ összesítése
- ▶ módosítása
- ▶ törlése
- ▶ tárolása

esetén **nem szerencsés** a redundancia

- ▶ Találjunk módot ezek kiküszöbölésére

- ▶ Megoldás: egyre szigorúbb formai szabályokat adunk meg a relációsémákra: **1NF – 2NF – 3NF ( – BCNF – 4NF)**

- ▶ A normalizálás tkp. mindig *célszerű*, de nem minden esetben *kötelező*

A ZH-ban és a dokumentációban viszont kötelező 😞

# I. NORMÁLFORMA (INF)

## Definíció.

Egy relációséma INF-ben van, ha az attribútumok értéktartománya csak egyszerű (atomi) adatokból áll.

- ▶ Nem lehet összetett attribútum v. egyéb nem atomi adat (pl. lista)
- ▶ Relációséma felírásakor kötelező az INF betartása
- ▶ Gyakorlatilag összetett *attribútumnál* az alábbi átalakítást jelenti:

Ügyfél(Azonosító, Név, Cím(Város, Utca, Házszám))



Ügyfél(Azonosító, Név, Város, Utca, Házszám)

- ▶ *Listáknál* gondoljunk vissza a többértékű attribútum leképezésére

## 2. NORMÁLFORMA (2NF)

### Definíció (teljes függés).

Legyen  $X, Y \subseteq A$ , és  $X \rightarrow Y$ . Azt mondjuk, hogy  $X$ -től **teljesen függ**  $Y$ , ha  $X$ -ből bármely attribútumot elhagyva a függőség már nem teljesül, vagyis bármely  $X_1 \subset X$  esetén  $X_1 \rightarrow Y$  már nem igaz.

### Definíció (2NF).

Egy relációséma 2NF-ben van, ha **minden** másodlagos attribútum teljesen függ **bármely** kulcstól.

### Az eddigiek képletesen.

Vannak másodlagos attribútumok, amik nem a teljes kulcstól, hanem annak csak egyes részeitől függenek.

## 2. NORMÁLFORMA (2NF)

### Következmény.

- (i) Az olyan sémák, ahol a kulcs egyszerű (csak egy attribútumból áll), mindig 2NF-ben vannak.
- (ii) Ha egy sémában nincsenek másodlagos attribútumok, akkor az a séma biztosan 2NF-ben van.

### Példa.

DOLGPROJ (Adószám, Név, Projektkód, Óra, Projektnév, Projekthely)

## 2. NORMÁLFORMA (2NF)

- ▶ A korábbi séma felett egy tábla lehetne pl. ilyen:

<u>Adószám</u>	Név	<u>Projekt kód</u>	Óra	Projekt név	Projekt hely
1111	Kovács	P2	4	Adatmodell	Veszprém
2222	Tóth	P1	6	Hardware	Budapest
4444	Kiss	P1	5	Hardware	Budapest
1111	Kovács	P1	2	Hardware	Budapest
1111	Kovács	P5	8	Teszt	Szeged

- ▶  $f_1: \text{Adószám} \rightarrow \text{Név}$
- ▶  $f_2: \text{Projekt kód} \rightarrow \{\text{Projekt név}, \text{Projekt hely}\}$
- ▶  $f_3: \{\text{Adószám}, \text{Projekt kód}\} \rightarrow \text{Óra}$

Név pl. csak Adószám-tól függ,  
így nem teljesen függ a kulcstól!

## 2. NORMÁLFORMÁRA HOZÁS

- ▶ Ha valamely  $K$  kulcsra  $L \subset K$  és  $L \rightarrow B$  (itt  $B$  az összes  $L$ -től függő másodlagos attribútum halmaza), akkor a sémát felbontjuk az  $L \rightarrow B$  függőség szerint.
- ▶ Legyen  $C = A - (LUB)$ , ekkor az  $R(A)$  sémát az  $R_1(CUL)$  és az  $R_2(LUB)$  sémákkal helyettesítjük.

### Példa.

Tétel (árúdkód, számla.sorszám, árunév, egységár, mennyiség)

1. 1NF? Teljesül ✓
2. 2NF? A kulcs összetett – vizsgáljuk meg alaposabban

## 2. NORMÁLFORMÁRA HOZÁS

Tétel (árukód, számla.sorszám, árunév, egységár, mennyiség)

Függőségek:

1.  $f_1: \{\text{számla.sorszám, árukód}\} \rightarrow \{\text{egységár, mennyiség}\}$
2.  $f_2: \{\text{árukód}\} \rightarrow \{\text{árunév}\}$

Az árunév másodlagos attribútum nem teljesen függ a kulcstól. → Nincs 2NF-ben X

## 2. NORMÁLFORMÁRA HOZÁS

Tétel (árukód, számla.sorszám, árunév, egységár, mennyiség)

**Megoldás.**

L                      B

A gondot okozó függés ( $f_2: \{\text{árukód}\} \rightarrow \{\text{árunév}\}$ ) mentén felbontjuk a sémát két másik sémára.

$C = A - (LUB), R(A) \rightarrow R_1(CUL), R_2(LUB)$

$L = \{\text{árukód}\}$

$B = \{\text{árunév}\}$

$LUB = \{\text{árukód}, \text{árunév}\}$

$C = A - (LUB) = \{\text{számla.sorszám}, \text{egységár}, \text{mennyiség}\}$

$CUL = \{\text{árukód}, \text{számla.sorszám}, \text{egységár}, \text{mennyiség}\}$

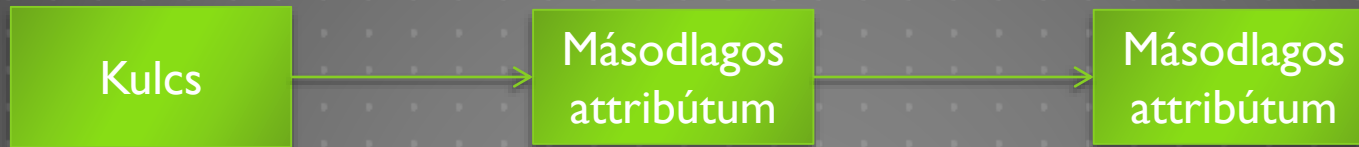
Tétel (árukód, számla.sorszám, egységár, mennyiség)

Áru (árukód, árunév)



# 3. NORMÁLFORMA (3NF)

Tranzitív függés.



Ellenkezője: közvetlen(ül) függés.

## 3NF.

Egy séma 3NF-ben van, ha minden másodlagos attribútum közvetlenül függ bármely kulcstól.

## Következmény.

- (i) Ha nincs a sémában másodlagos attribútum, akkor biztosan 3NF-ben van.

# 3. NORMÁLFORMÁRA HOZÁS

- ▶ Megoldás: ismételten függőség szerinti felbontást végzünk

## Példa.

Számla (sorszám, dátum, vevőkód, vevőnév, vevőcím)

1. 1NF? Teljesül ✓
2. 2NF? Egyszerű kulcs – teljesül ✓
3. 3NF? Vizsgáljuk meg a függőségeket!

$\{\text{sorszám}\} \rightarrow \{\text{vevőkód}\} \rightarrow \{\text{vevőnév, vevőcím}\}$

(Adott sorszámú számlát adott vevőnek állítunk ki, és a vevőkód alapján kiderül a vevő neve és címe is.)

# 3. NORMÁLFORMÁRA HOZÁS

**Megoldás.**

Számla (sorszám, dátum, vevőkód)

Vevő (vevőkód, vevőnév, vevőcím)

Vagyis, felbontást végeztünk a  $\{\text{sorszám}\} \rightarrow \{\text{vevőkód}\} \rightarrow \{\text{vevőnév}, \text{vevőcím}\}$  függőség szerint.



# FELADATOK

5.3 Hozzuk 1,2,3NF-re az alábbi relációsémát a kulcs bejelölése után:

Hallgató(éha, név, város, irányítószám, utca, házszám, szak, kar)

5.4 Hozzuk 1,2,3NF-re az alábbi relációsémát a kulcs bejelölése után:

Áram( ünév, vóra\_ száma, szavatosság, mérés\_ kezd, mérés\_ vége, ücím)

ahol az ücím attribútum összetett attribútum (város, utca, házszám, irsz).

# FORRÁSOK

- ▶ [1] Dr. Katona Endre: Adatbázisok
- ▶ [2] Németh Gábor: Kidolgozott példák