

Adatbázisok 1.

Relációs adatbázis tervezés – 1. rész

Funkcionális függőségek

Felbontások

Normálformák

A reláció adatmodell tervezése

	A	B	C	D	E	F	G
1	teázó_város ▾	teázó_név ▾	tulaj ▾	tea ▾	gyártó ▾	ár_USD ▾	ár_HUF ▾
2	New York	Green Teahouse	Joe	Brisk	Lipton	3	1200
3	New York	Green Teahouse	Joe	Pyramid	Lipton	2,5	1000
4	Washington	Green Teahouse	Sue	Brisk	Lipton	3,5	1400
5	London	Frank's Twinings	Frank	Brisk	Lipton	3	1200
6	London	Frank's Twinings	Frank	Earl Grey	Tetley	3,5	1400
7

A reláció adatmodell tervezése

	A	B	C	D	E	F	G
1	teázó_város ▾	teázó_név ▾	tulaj ▾	tea ▾	gyártó ▾	ár_USD ▾	ár_HUF ▾
2	New York	Green Teahouse	Jack	Brisk	Lipton	3	1200
3	New York	Green Teahouse	Joe	Pyramid	Lipton	2,5	1000
4	Washington	Green Teahouse	Sue	Brisk	Lipton	3,5	1400
5	London	Frank's Twinings	Frank	Brisk	Lipton	3	1200
6	London	Frank's Twinings	Frank	Earl Grey	Tetley	3,5	1400
7

A reláció adatmodell tervezése

	A	B	C	D	E	F	G
1	teázó_város ▾	teázó_név ▾	tulaj ▾	tea ▾	gyártó ▾	ár_USD ▾	ár_HUF ▾
2	New York	Green Teahouse	Joe	Brisk	Lipton	4	1200
3	New York	Green Teahouse	Joe	Pyramid	Lipton	2,5	1000
4	Washington	Green Teahouse	Sue	Brisk	Lipton	3,5	1400
5	London	Frank's Twinings	Frank	Brisk	Lipton	3	1200
6	London	Frank's Twinings	Frank	Earl Grey	Tetley	3,5	1400
7

A relációs adatmodell tervezése

- Megtehetnénk, hogy valamennyi adatot egyetlen relációba tennénk:
 - Felhasználó szempontjából kényelmes
 - Sok „felesleges” adat van benne → nem jó tárolási hatékonyság, az adatbázist ellentmondásossá teheti

A relációs adatmodell tervezése

- Előzők miatt → valahogy a felhasználási eset fogalmait, kapcsolatait **modellezni kell** (pl. Egyed/Kapcsolat modellel, ld. későbbi előadáson)
- De még így is előfordulhat, hogy nem lesz elég „ügyes” a modellezés
→ még mindig lehetnek felesleges adatok konkrét előfordulásnál

A relációs adatmodell tervezése

	A	B	C	D	E	F	G
1	teázó_város	teázó_név	tulaj	tea	gyártó	ár_USD	ár_HUF
2	New York	Green Teahouse	Joe	Brisk	Lipton	3	1200
3	New York	Green Teahouse	Joe	Pyramid	Lipton	2,5	1000
4	Washington	Green Teahouse	Sue	Brisk	Lipton	3,5	1400
5	London	Green Teahouse	Frank	Brisk	Lipton	2	1200

Felhasználó (+)

- Fontos fogalom a redundancia (*redundancy*):
- Adott egy $R(A_1, \dots, A_n)$ reláció. Ha valamely A_i attribútum értékét ki tudjuk az $\{A_j \mid j \neq i\}$ attribútumok értékeiből „következtetni” \rightarrow a relációt redundánsnak (*redundant*) nevezzük.
- A „következtetési szabályok” az ún. Armstrong-axiómák lesznek (ld. alább)

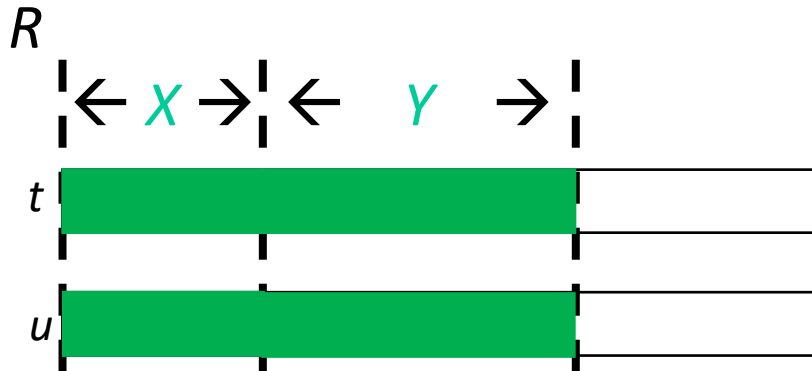
Funkcionális függőségek (*functional dependencies*)

	A	B	C	D	E	F	G
1	teázó_város	teázó_név	tulaj	tea	gyártó	ár_USD	ár_HUF
2	New York	Green Teahouse	Joe	Brisk	Lipton	3	1200
3	New York	Green Teahouse	Joe	Pyramid	Lipton	2,5	1000
4	Washington	Green Teahouse	Sue	Brisk	Lipton	3,5	1400
5	London	Frank's Teahouse	Frank	Brisk	Lipton	2	1200

- Másképp:
- teázó_város és teázó_név mezők egyezőségéből következik a tulaj mezők egyezősége
- Létezik egy függvény (funkció), amely
(teázó_város, teázó_név) → tulaj

Funkcionális függőségek (*functional dependencies*)

- $X \rightarrow Y$ egy R relációra vonatkozó megszorítás, miszerint ha két sor megegyezik X összes attribútumán, Y attribútumain is meg kell, hogy egyezzenek.
 - Jelölés: X, Y, Z, \dots attribútum halmazokat; A, B, C, \dots attribútumokat jelöl.
 - Jelölés: $\{A, B, C\}$ attribútum halmaz helyett ABC -t írunk.



Funkcionális függőségek – dekompozíció

	A	B	C	D
1	teázó_város ▾	teázó_név ▾	tea ▾	ár_USD ▾
2	New York	Green Teahouse	Brisk	3
3	New York	Green Teahouse	Pyramid	2,5
4	Washington	Green Teahouse	Brisk	3,5
5	London	Frank's Twinings	Brisk	3
6	London	Frank's Twinings	Earl Grey	3,5
7

Felszolgál | Teázók | Teák | (+)

	A	B	C
1	teázó_város ▾	teázó_név ▾	tulaj ▾
2	New York	Green Teahouse	Joe
3	Washington	Green Teahouse	Sue
4	London	Frank's Twinings	Frank
5

Felszolgál | Teázók | Teák | (+)

	A	B
1	tea ▾	gyártó ▾
2	Brisk	Lipton
3	Pyramid	Lipton
4	Earl Grey	Tetley
5

Felszolgál | Teázók | Teák | (+)

Jobboldalak szétvágása (ff)

- $X \rightarrow A_1 A_2 \dots A_n$ akkor és csak akkor teljesül R relációra, ha $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$ is teljesül R -en.
- Példa: $A \rightarrow BC$ ekvivalens $A \rightarrow B$ és $A \rightarrow C$ függőségek kettőssével.
- Baloldalak szétvágására nincs általános szabály.
- Általában FF-k jobboldalán egyetlen attribútum szerepel majd.

Példa: FF

Főnökök(név, cím, kedveltTeák, gyártó, kedvencTea)

- FF-k, amelyek vszleg teljesülnek:
 1. név -> cím kedvencTea
 - Ez az FF ugyanaz, mint név -> cím és név -> kedvencTea.
 2. kedveltTeák -> gyártó.

Példa: egy lehetséges előfordulás

név	cím	kedveltTeák	gyártó	kedvencTea
Janeway	Voyager	Brisk	Lipton	E. G.
Janeway	Voyager	E. G.	Tetley	E. G.
Spock	Enterprise	Brisk	Lipton	Brisk

Mert név -> cím

Mert név -> kedvencTea

Mert kedveltTeák -> gyártó

Relációk kulcsai

- K *szuperkulcs* R relációra, ha K funkcionálisan meghatározza R attribútumait.
- K *kulcs* R -en, ha K szuperkulcs, de egyetlen valódi részhalmaza sem szuperkulcs.

Példa: superkulcs

Főnökök(név, cím, kedveltTeák, gyártó, kedvencTea)

□ {név, kedveltTeák} superkulcs, hiszen a két attribútum meghatározza funkcionálisan a maradék attribútumokat.

➤ név -> cím kedvencTea

➤ kedveltTeák -> gyártó

Példa: kulcs

- {név, kedveltTeák} kulcs, hiszen sem {név}, sem {kedveltTeák} nem superkulcs.
 - név -> gyártó; kedveltTeák -> cím nem teljesülnek.
- Az előbbin kívül nincs több kulcs, de számos superkulcs megadható még.
 - Minden olyan halmaz, amit tartalmazza {név, kedveltTeák}-t.

Kis kombinatorika

- **Feladat:** R relációnak legyenek A_1, \dots, A_n az attribútumai. Adjuk meg n függvényeként, hogy R -nek hány superkulcsa van, ha
 - (a) csak A_1 kulcs,
 - (b) A_1 és A_2 kulcsok,
 - (c) $\{A_1, A_2\}, \{A_3, A_4\}$ kulcsok,
 - (d) $\{A_1, A_2\}, \{A_1, A_3\}$ kulcsok.

Hogyan kaphatjuk meg a kulcsokat?

1. Szimplán megadunk egy K kulcsot, mert a specifikáció alapján eldönthető.
 - Az FF-k $K \rightarrow A$ alakúak, ahol A „végigmegy” az összes attribútumon
2. Vagy: megadjuk az FF-eket, és ezekből következtetjük ki a kulcsokat.

Még egy természetesen adódó FF

- **Példa:** az „ugyanabban az időben nem lehet két előadás ugyanabban a teremben” lefordítva:
idő terem -> előadás.

FF-k kikövetkeztetése (*inferring*)

- Legyenek $X_1 \rightarrow A_1, X_2 \rightarrow A_2, \dots, X_n \rightarrow A_n$ adott FF-ek, szeretnénk tudni, hogy $Y \rightarrow B$ teljesül-e olyan relációkra, amire az előbbi FF-ek teljesülnek.
 - Példa: $A \rightarrow B$ és $B \rightarrow C$ teljesülése esetén $A \rightarrow C$ biztosan teljesül.
- Ez az adatbázis sémájának megtervezésekor lesz majd fontos.

Armstrong-axiómák I.

- (A1) **Reflexivitás** (*reflexivity*): ha $Y \subseteq X \subseteq R$, akkor $X \rightarrow Y$. Az ilyen függőségeket **triviális** függőségeknek nevezzük.
- (A2) **Bővítés** (*augmentation*): ha $X \rightarrow Y$ teljesül, akkor tetszőleges $Z \subseteq R$ -ra $XZ \rightarrow YZ$ teljesül.
- (A3) **Tranzitivitás** (*transitivity*): ha $X \rightarrow Y$ és $Y \rightarrow Z$, akkor $X \rightarrow Z$.
- Példák a *személy* (*sz_ig_száma*, *TAJ*, *név*, *anyja_neve*, *születés*, *kor*, *fizetés*) tábla esetén:
 - (A1) (*név*, *születés*) \rightarrow *név*
 - (A2) *születés* \rightarrow *kor*, akkor (*születés*, *név*) \rightarrow (*kor*, *név*)
 - (A3) *TAJ* \rightarrow *születés*, *születés* \rightarrow *kor*, akkor *TAJ* \rightarrow *kor*.

Példa levezetésre

- Legyen $R = ABCD$ és $F = \{ A \rightarrow C, B \rightarrow D \}$:
 1. $A \rightarrow C$ adott.
 2. $AB \rightarrow ABC$ (A2) alapján.
 3. $B \rightarrow D$ adott.
 4. $ABC \rightarrow ABCD$ (A2) alapján.
 5. $AB \rightarrow ABCD$ (A3) alapján 2-ből és 4-ből.
- **Példa:** bizonyítsuk be levezetéssel, hogy $\{ X \rightarrow Y, XY \rightarrow Z \}$ -ből következik $\{ X \rightarrow Z \}$.

Újabb feladat

- **Feladat:** mutassuk meg, hogy az alábbiak nem érvényes szabályok funkcionális függőségekre:
 - ha $A \rightarrow B$, akkor $B \rightarrow A$,
 - ha $AB \rightarrow C$ és $A \rightarrow C$, akkor $B \rightarrow C$,
 - ha $AB \rightarrow C$, akkor $A \rightarrow C$ vagy $B \rightarrow C$.