

Adatbázis-kezelő rendszerek I.



NORMALIZÁCIÓ

Redundáns adattárolás és anomáliák



- **Redundancia:** többszörös felesleges adattárolás
- Redundáns információ okozhat:
 - Felesleges helyfoglalás
 - **Anomáliák:**
 - ✦ **Beszúrási anomália**
 - ✦ **Törlési anomália**
 - ✦ **Módosítási anomália**
- **Adatbázis konzisztencia:** Az adatbázis csak valós adatokat tartalmazhat.
 - A redundancia adat **inkonzisztenciát** okozhat.
 - Az anomáliák *inkonzisztens adatbázis állapothoz* vezethetnek.

Anomáliák



- **Beszúrási anomália:**
 - Akkor következik be, ha bizonyos attribútumokat nem tudunk beilleszteni az adatbázisba más attribútumok hiányában.
- **Módosítási anomália:**
 - Akkor beszélhetünk róla, ha adatmódosítás során a redundánsan tárolt adatok esetében nem minden duplumot módosítunk, s ezáltal ellentmondó információ keletkezik.
- **Törlési anomália:**
 - Bizonyos adatok (rekordok) törlése olyan adatok elvesztését eredményezi, melyet nem szerettünk volna törölni.

Két relációs séma, amely felesleges helyfoglalást és anomáliákat hordoz magában



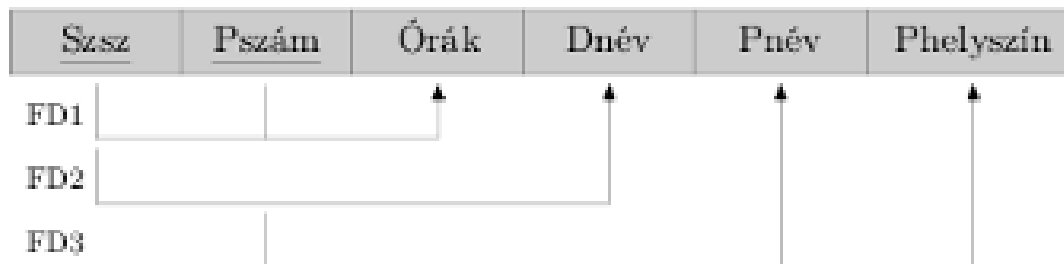
(a)

DOLG_OSZT



(b)

DOLG_PROJ



DOLG_OSZT és DOLG_PROJ relációk lehetséges adatbázis állapota

DOLG_OSZT

Dnév	Szzs	Szdátum	Lakcím	Oszám	redundancia	
					Onév	Ovez_szzs
Kovács László	1 650109 0812	1965. január 9.	4033 Debrecen	5	Kutatás	2 551208 2219
Szabó Mária	2 551208 2219	1955. december 8.	1097 Budapest	5	Kutatás	2 551208 2219
Kiss István	1 680119 6749	1968. január 19.	1172 Budapest	4	Humán erőforrás	2 690329 1099
Takács József	1 410620 4902	1941. június 20.	4027 Debrecen	4	Humán erőforrás	2 690329 1099
Horváth Erzsébet	2 620915 3134	1962. szeptember 15.	1092 Budapest	5	Kutatás	2 551208 2219
Tóth János	1 720731 2985	1972. július 31.	6726 Szeged	5	Kutatás	2 551208 2219
Fazekas Ilona	2 690329 1099	1969. március 29.	3535 Miskolc	4	Humán erőforrás	2 690329 1099
Nagy Zoltán	1 371110 4519	1937. november 10.	1061 Budapest	1	Központ	1 371110 4519

DOLG_PROJ

Szzs	Pszám	Órák	Dnév	redundancia	
				Pnév	Phelyszín
1 650109 0812	1	32.5	Kovács László	X termék	Vác
1 650109 0812	2	7.5	Kovács László	Y termék	Tiszafüred
2 620915 3134	3	40.0	Horváth Erzsébet	Z termék	Budapest
1 720731 2985	1	20.0	Tóth János	X termék	Vác
1 720731 2985	2	20.0	Tóth János	Y termék	Tiszafüred
2 551208 2219	2	10.0	Szabó Mária	Y termék	Tiszafüred
2 551208 2219	3	10.0	Szabó Mária	Z termék	Budapest
2 551208 2219	10	10.0	Szabó Mária	Komputerizáció	Kecskemét
2 551208 2219	20	10.0	Szabó Mária	Reorganizáció	Budapest
1 680119 6749	30	30.0	Kiss István	Új fejlesztések	Kecskemét
1 680119 6749	10	10.0	Kiss István	Komputerizáció	Kecskemét
2 690329 1099	10	35.0	Fazekas Ilona	Komputerizáció	Kecskemét
2 690329 1099	30	5.0	Fazekas Ilona	Új fejlesztések	Kecskemét
1 410620 4902	30	20.0	Takács József	Új fejlesztések	Kecskemét
1 410620 4902	20	15.0	Takács József	Reorganizáció	Budapest
1 371110 4519	20	NULL	Nagy Zoltán	Reorganizáció	Budapest

Módosítási anomália – Példa



- **Módosítási anomália:**

- Példa1: Tekintsük a következő relációt:

DOLG_PROJ(Szsz, Pszám, Órák, Dnév, Pnév, Phelyszín)

Amennyiben a 1-es projekt helyszínét “Vác”-ról áthelyezzük “Budapest”-re, akkor ezt a módosítást minden 1-es projekten dolgozó alkalmazott esetében el kell végezni (lehet akár több 100 is).

- Példa2: A DOLG_OSZT relációban amennyiben megváltoztatjuk az egyik részleg attribútumát (pld: az osztály vezetőjét) módosítanunk kell az összes olyan alkalmazott rekordját, akik azon a részlegen dolgoznak.
- Máskülönben az adatbázis **inkonzisztensé** válik.

Beszűrési anomália – Példa



- Tekintsük a következő relációt:

DOLG_PROJ(Szsz, Pszám, Órák, Dnév, Pnév, Phelyszín)

- **Beszűrési anomália:**

- Nem tudunk projektet felvenni, ha nem rendelünk hozzá egy alkalmazottat.
- Fordítva: Nem tudunk felvenni alkalmazottat, ha nincs hozzárendelve projekthez.

Törlési anomália – Példa



- Tekintsük a következő relációt:
DOLG_PROJ(Szsz, Pszám, Órák, Dnév, Pnév, Phelyszín)
- **Törlési anomália:**
 - Amikor egy projektet törlünk, az eredményez(het)i az összes alkalmazott törlését is, aki az adott projekten dolgozik.
 - Másrésről, ha egy alkalmazott az egyetlen alkalmazott a projekten, az alkalmazott törlésével a szóban forgó projekt törlése is bekövetkezik.

Funkcionális függőségek



Funkcionális függőség (FD)



Definíció:

Az $X \rightarrow Y$ **funkcionális függőség** olyan korlátozást definiál az R -beli X és Y attribútumhalmazok között, amelynek minden érvényes R -beli r állapotban teljesülnie kell.

A korlátozás: bármely érvényes $r(R)$ állapot esetén tetszőlegesen választott t_1 és t_2 rekordokra, amennyiben $t_1[X] = t_2[X]$ teljesül, akkor $t_1[Y] = t_2[Y]$ -nek is teljesülnie kell.

Például:

Legyenek A és B egy R reláció attribútumai. B **funkcionálisan függ** A -tól (jelölése $A \rightarrow B$), Ha R -ben A minden egyes értékéhez B pontosan egy értéke tartozik.

✦ A és B lehetnek akár attribútumhalmazok is.

- $X \rightarrow Y$, elnevezések:
 - Y **funkcionálisan függ** X -től
 - Az X komponens **funkcionálisan meghatározza** az Y komponenst.
 - X : **bal oldala** a funkcionális függőségnek
 - Y : **jobb oldala** a funkcionális függőségnek

Funkcionális függőség (folyt.)



- Ha $X \rightarrow Y$ igaz R -ben, akkor $Y \rightarrow X$ –ről nem mondtunk semmit, sem azt hogy igaz, sem azt hogy nem.
- Ha K egy kulcs az R relációban, akkor a K funkcionálisan meghatároz minden attribútumot az R relációban.
 - Mivel nem lehet 2 olyan különböző sorunk, ahol $t1[K]=t2[K]$ igaz lenne.

DOLGOZÓ

Vnév	Knév	Szsz	Szdátum	Lakcím	Nem	Fizetés	Főnök_szsz	Osz
Kovács	László	1 650109 0812	1965. január 9.	4033 Debrecen	F	390000	2 551208 2219	5
Szabó	Mária	2 551208 2219	1955. december 8.	1097 Budapest	N	520000	1 371110 4519	5
Kiss	István	1 680119 6749	1968. január 19.	1172 Budapest	F	325000	1 410620 4902	4
Takács	József	1 410620 4902	1941. június 20.	4027 Debrecen	F	559000	1 371110 4519	4
Horváth	Erzsébet	2 620915 3134	1962. szeptember 15.	1092 Budapest	N	494000	2 551208 2219	5
Tóth	János	1 720731 2985	1972. július 31.	6726 Szeged	F	325000	2 551208 2219	5
Fazekas	Ilona	2 690329 1099	1969. március 29.	3535 Miskolc	N	325000	1 410620 4902	4
Nagy	Zoltán	1 371110 4519	1937. november 10.	1061 Budapest	F	715000	NULL	1

OSZTÁLY

Onév	Oszám	Vez_szsz	Vez_kezdő_dátum
Kutatás	5	2 551208 2219	1988. május 22.
Humán erőforrás	4	2 690329 1099	1995. január 1.
Központ	1	1 371110 4519	1981. június 19.

DOLGOZIK_RAJTA

Dszsz	Psz	Órák
1 650109 0812	1	32.5
1 650109 0812	2	7.5
2 620915 3134	3	40.0
1 720731 2985	1	20.0
1 720731 2985	2	20.0
2 551208 2219	2	10.0
2 551208 2219	3	10.0
2 551208 2219	10	10.0
2 551208 2219	20	10.0
1 680119 6749	30	30.0
1 680119 6749	10	10.0
2 690329 1099	10	35.0
2 690329 1099	30	5.0
1 410620 4902	30	20.0
1 410620 4902	20	15.0
1 371110 4519	20	NULL

OSZT_HELYSZÍNEK

Oszám	Ohelyszín
1	Budapest
4	Kecskemét
5	Vác
5	Tiszafüred
5	Budapest

PROJEKT

Pnév	Pszám	Phelyszín	Osz
X termék	1	Vác	5
Y termék	2	Tiszafüred	5
Z termék	3	Budapest	5
Komputerizáció	10	Kecskemét	4
Reorganizáció	20	Budapest	1
Új fejlesztések	30	Kecskemét	4

HOZZÁTARTOZÓ

Dszsz	Hozzá tartozó_név	Nem	Szdátum	Kapcsolat
2 551208 2219	Anna	N	1986. április 5.	lánya
2 551208 2219	Bence	F	1983. október 25.	fia
2 551208 2219	Máté	F	1958. május 3.	házastársa
1 410620 4902	Viktória	N	1942. február 28.	házastársa
1 650109 0812	Balázs	F	1988. január 4.	fia
1 650109 0812	Anna	N	1988. december 30.	lánya
1 650109 0812	Réka	N	1967. május 5.	házastársa

Példa – VÁLLALAT

A VÁLLALAT
relációs adatbázis séma
egy lehetséges állapota.

Példák funkcionális függőségekre



- A *Szsz* értéke meghatározza az alkalmazott vezetéknévét:
 - *Szsz* -> *Vnév*
- A projekt száma meghatározza a projekt nevét és helyét:
 - *Pszám* -> *Pnév*, *Phelyszín*
- A dolgozó személyi száma (*Szsz*) és a projekt száma meghatározza, hogy egy dolgozó az adott projekten hetente hány órát dolgozik:
 - $\{Dszsz, Pszám\}$ -> *Órák*

A funkcionális függőségek következményei



- Adott F funkcionális függőségek halmaza. Ebből a halmazból **levezethetők** további funkcionális függőségek is, amelyek mindakkor igazak lesznek, ha az F -beli függőségek igazak.
- **Armstrong axiómák:**
 - A1. (**Reflexivitás**) Ha $Y \subseteq X \subseteq R$, akkor $X \rightarrow Y$
 - A2. (**Bővítés**) Ha $X \rightarrow Y$ teljesül, akkor tetszőleges $Z \subseteq R$ -ra $XZ \rightarrow YZ$ is teljesül
 - ✦ (megj: XZ az $X \cup Z$ -t jelöli)
 - A3. (**Tranzitivitás**) Ha $X \rightarrow Y$ és $Y \rightarrow Z$ teljesül, akkor $X \rightarrow Z$ is igaz.
- Néhány további hasznos következmény:
 - **Szétvágási szabály:** Ha $X \rightarrow YZ$ igaz, akkor $X \rightarrow Y$ és $X \rightarrow Z$ is teljesülnek
 - **Egyesítési szabály:** Ha $X \rightarrow Y$ és $X \rightarrow Z$ igaz, akkor $X \rightarrow YZ$ is teljesül
 - **Pszédotranzitivitás:** Ha $X \rightarrow Y$ és $WY \rightarrow Z$ igaz, akkor $WX \rightarrow Z$ is igaz
- Az utolsó 3 következmény (mint egyéb következmények is) levezethetők az A1, A2, és A3 axiómák segítségével (teljességi tétel).

Normalizáció



Normalizáció



- **Normalizáció:**

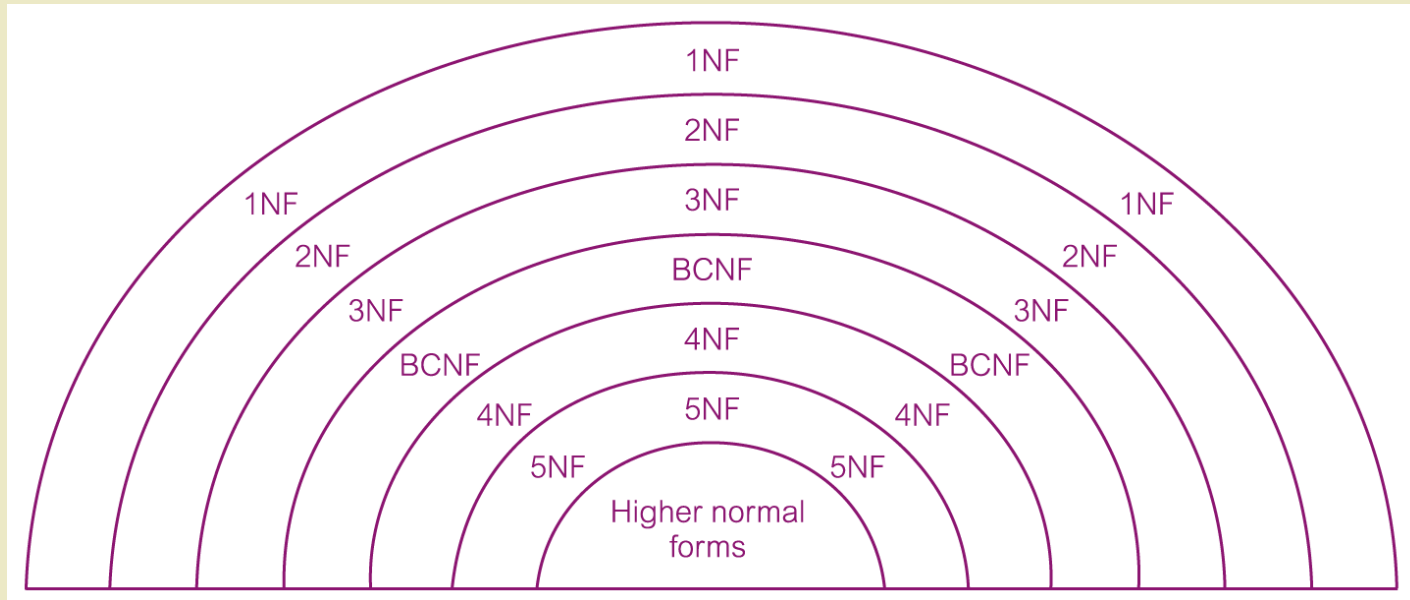
- Relációs adatbázis-tervezési folyamat (módszertan).
- A folyamat során a nem megfelelő, „rossz” relációkat attribútumaik mentén kisebb relációkra osztjuk.
- A **normalizáció** egy felülről lefelé (top-down) haladó dekompozíciós eljárás, amikor azon relációkat, amelyek nem elégítenek ki bizonyos feltételeket – a normál formák tesztje – olyan kisebb relációkká bontjuk szét, amelyek kielégítik ezen feltételeket.
- A normalizáció végrehajtásához **azonosítani kell azokat a funkcionális függőségeket** amelyek **minden érvényes reláció előfordulásban igazak**.
- Szekvenciális lépések sorozata, ahol minden egyes lépés egy meghatározott normál formához vezet.

Normál formák (NF-k)



- **Normál forma:**
 - Kulcsokat és funkcionális függőségeket alkalmazó szabályok, melyek segítségével meghatározható, hogy a relációs séma milyen normál formában van.
 - *Def:* **Egy reláció normál formája** az a legmagasabb normál forma, melynek feltételeit kielégíti .
- UNF (0NF):
 - A tábla, amely egy vagy több ismétlődő csoportot tartalmaz
- 1NF:
 - A reláció formális definícióján alapul.
- 2NF, 3NF, BCNF
 - A kulcsok és funkcionális függőségek alapján.
- 4NF
 - A többértékű függőségek alapján.
- 5NF
 - Az összekapcsolási függőségek alapján.

Normál formák (folyt.)



A normál formák kapcsolata:

Látható, hogy azok a relációk, ami 2NF-ben vannak 1NF-ben is vannak, azok a relációk amik 3NF-ben vannak, 2NF-ben is vannak , és így tovább.

Nulladik normál forma (0NF, UNF)

– Normalizálatlan relációs séma



- A normalizáció előtt:
 - Vegyük az összes mezőt, amelyet az adatbázisnak tartalmaznia kell és tekintsünk erre mint egy nagy táblázatra.
- **Nulladik normál forma (UNF):** Egy táblázat, amely egy vagy több ismétlődő csoportot tartalmaz.
- **Ismétlődő csoport:** A táblázaton belül egy ismétlődő attribútum, vagy attribútum csoport, amely egy kulcsjelölt esetében több értéket is felvehet.

Első normál forma (1NF)



- **1 NF:**
 - kielégíti a reláció matematikai definícióját ÉS
 - meghatároztuk az elsődleges kulcsot!
- Nem megengedettek:
 - összetett attribútum
 - többértékű attribútum
 - beágyazott reláció: olyan attribútum, amelynek értéke egy rekord esetében nem atomi

Normalizáció 1NF-be



Feltételezzük, hogy egyazon osztály a vállalat több telephelyén is jelen lehet.

(a)

OSZTÁLY

Onév	<u>Oszám</u>	Ovez_szzsz	Ohelyszínek

(b)

OSZTÁLY

Onév	<u>Oszám</u>	Ovez_szzsz	Ohelyszínek
Kutatás	5	2 551208 2219	{ Vác, Tiszafüred, Budapest }
Humán erőforrás	4	2 690329 1099	{ Kecskemét }
Központ	1	1 371110 4518	{ Budapest }

(c)

OSZTÁLY

Onév	<u>Oszám</u>	Ovez_szzsz	<u>Ohelyszín</u>
Kutatás	5	2 551208 2219	Vác
Kutatás	5	2 551208 2219	Tiszafüred
Kutatás	5	2 551208 2219	Budapest
Humán erőforrás	4	2 690329 1099	Kecskemét
Központ	1	1 371110 4518	Budapest

Elsődleges kulcs!

Beágyazott relációk normalizációja 1NF-be

(a)

DOLG_PROJ

Projektek

Szsz	Dnév	Pszám	Órák
------	------	-------	------



1NF:

(b)

DOLG_PROJ

Szsz	Dnév	Pszám	Órák
1 650109 0812	Kovács László	1	32.5
		2	7.5
2 620915 3134	Horváth Erzsébet	3	40.0
1 720731 2985	Tóth János	1	20.0
		2	20.0
2 551208 2219	Szabó Mária	2	10.0
		3	10.0
		10	10.0
		20	10.0
1 680119 6749	Kiss István	30	30.0
		10	10.0
2 690329 1099	Fazekas Ilona	10	35.0
		30	5.0
1 410620 4902	Takács József	30	20.0
		20	15.0
1 371110 4519	Nagy Zoltán	20	NULL

DOLG_PROJ(Szsz, Dnév, Pszám, Órák)

Második normál forma(2NF)



- Az elsődleges kulcs és a funkcionális függőség definíciója alapján
- Definíciók:
 - **Elsődleges attribútum:** egy attribútum, amely az elsődleges kulcs része
 - **Teljes funkcionális függőség:** olyan $Y \rightarrow Z$ függőség, ahol ha az Y attribútum halmazból bármely attribútumot(ka)t elveszünk, akkor a fennmaradó attribútum halmaz már nem határozza meg funkcionálisan Z -t.
 - **Részleges funkcionális függőség:** Egy nem elsődleges attribútum az elsődleges kulcsnak csak egy részétől függ funkcionálisan.
- Példák:
 - $\{Szs z, Pszám\} \rightarrow Órák$ egy teljes FD, mivel sem $Szs z \rightarrow Órák$, sem $Pszám \rightarrow Órák$ nem igaz
 - $\{Szs z, Pszám\} \rightarrow Vnév$ nem teljes FD (hanem részleges), mivel $Szs z \rightarrow Vnév$ igaz

Második normál forma(2NF) (folyt.)



Definíció:

Egy R relációs séma **2NF-ban van**, ha minden nem elsődleges attribútuma *teljesen funkcionálisan függ* R elsődleges kulcsától.

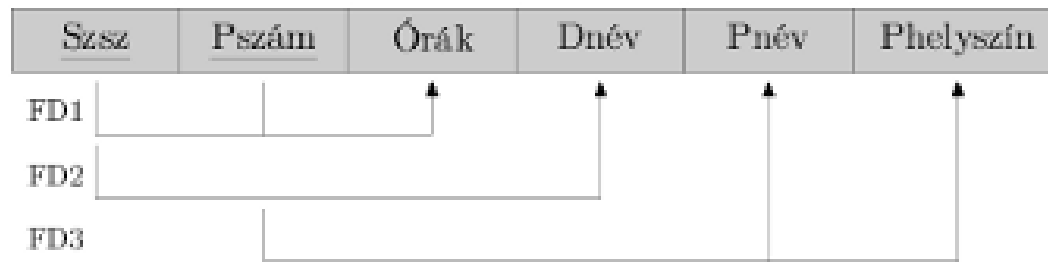
- R 2NF-be hozható a 2NF dekompozíciós eljárás segítségével.

Normalizáció 2NF-be



(a)

DOLG_PROJ



2NF normalizálás

DP1



DP2



DP3



Harmadik normál forma (3NF)



- Definíció:
 - **Tranzitív funkcionális függőség:** az $X \rightarrow Z$ funkcionális függőség tranzitív, ha létezik olyan Y attribútumhalmaz, amelyre $X \rightarrow Y$ és $Y \rightarrow Z$ is teljesül
- Példák:
 - $Szsz \rightarrow Főnök_szsz$ egy **tranzitív** FD
 - ✦ Mivel $Szsz \rightarrow Oszám$ és $Oszám \rightarrow Főnök_szsz$ igaz
 - $Szsz \rightarrow Vnév$ **nem tranzitív** FD
 - ✦ Mivel nincs olyan X attribútumhalmaz, ahol: $Szsz \rightarrow X$ és $X \rightarrow Vnév$ igaz lenne

Harmadik normál forma (3NF) (folyt.)



Definition:

Egy R séma **3NF-ben van**, ha 2NF-ben van és nincs olyan nem elsődleges attribútuma, ami tranzitív módon függne az elsődleges kulcstól.

- R 3NF-be hozható a 3NF dekompozíciós eljárás segítségével
- Megj:
 - Az $X \rightarrow Y$ és $Y \rightarrow Z$ esetében, ahol X elsődleges kulcs, a tranzitív függőséggel problémájával csak akkor foglalkozunk, ha Y nem kulcsjelölt.
 - Ha Y kulcsjelölt, a tranzitív függőség nem okoz problémát.
 - Pl., tekintsük az DOLG(Szsz, Dolgozo_azon, Fizetés) relációt.
 - ✦ Itt, $Szsz \rightarrow Dolgozo_azon \rightarrow Fizetés$ igaz és $Dolgozo_azon$ kulcsjelölt.

Normalizáció 3NF-be



(b)

DOLG_OSZT



3NF normalizálás

DO1



DO2



Általános definíciók



- Általános definíciók:
 - Első normál forma (1NF)
 - ✦ A táblázat első normál formában van, ha:
 - Van elsődleges kulcsa.
 - Egyetlen attribútum sem vesz fel egy cellában több értéket.
 - A nem kulcs attribútumok függnék az elsődleges kulcstól.
 - Második normál forma (2NF)
 - ✦ A reláció második normál formában van, ha első normál formában van és minden nem elsődleges kulcs attribútuma teljesen funkcionálisan függ bármely kulcsjelölttől.
 - Harmadik normál forma (3NF)
 - ✦ A reláció harmadik normál formában van, ha második normál formában van és nincs olyan nem elsődleges kulcs attribútuma, amely tranzitív módon függne bármely kulcsjelölttől.

Boyce–Codd normál forma (BCNF)



Definíció:

Az R relációs séma **Boyce-Codd normál formában** van (BCNF), ha valahányszor egy $X \rightarrow A$ funkcionális függőség teljesül R -ben, akkor X az R relációs séma **szuperkulcsa**.

- Különbség a 3NF és BCNF között:
 - A 3NF megenged olyan $A \rightarrow B$ funkcionális függőséget, ahol B elsődleges kulcs attribútum és A nem kulcsjelölt.
 - Ugyanakkor a BCNF ragaszkodik hozzá, hogy a fenti funkcionális függőség esetében az A -nak kulcsjelöltnek kell lennie.
- Minden BCNF-ben lévő reláció egyben 3NF-ben is van. Ugyanakkor a 3NF-ben lévő reláció nincs feltétlenül BCNF-ben.
 - A BCNF egy szigorúbb követelmény.

Egy reláció, ami 3NF-ben van, de nincs BCNF-ben



Napi teniszpálya bérlek.

A teniszklubban két pálya van:

Court 1: salakos pálya

Court 2: füves pálya

4 árkategória:

- SAVER: a Court 1- re tagoknak
- STANDARD: Court 1-re nem tagoknak
- PREMIUM-A: Court 2-re tagoknak
- PREMIUM-B: Court 2-re nem tagoknak

Today's Court Bookings

<u>Court</u>	<u>Start Time</u>	<u>End Time</u>	<u>Rate Type</u>
1	09:30	10:30	SAVER
1	11:00	12:00	SAVER
1	14:00	15:30	STANDARD
2	10:00	11:30	PREMIUM-B
2	11:30	13:30	PREMIUM-B
2	15:00	16:30	PREMIUM-A

Egy reláció, ami 3NF-ben van, de nincs BCNF-ben

Szuperkulcsok:

$S_1 = \{\text{Court, Start Time}\}$

$S_2 = \{\text{Court, End Time}\}$

$S_3 = \{\text{Rate Type, Start Time}\}$

$S_4 = \{\text{Rate Type, End Time}\}$

$S_5 = \{\text{Court, Start Time, End Time}\}$

$S_6 = \{\text{Rate Type, Start Time, End Time}\}$

$S_7 = \{\text{Court, Rate Type, Start Time}\}$

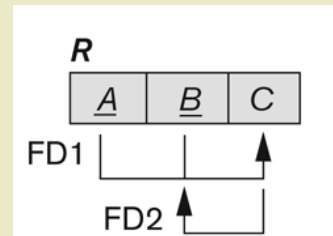
$S_8 = \{\text{Court, Rate Type, End Time}\}$

$S_T = \{\text{Court, Rate Type, Start Time, End Time}\}$ triviális szuperkulcs.

Today's Court Bookings

<u>Court</u>	<u>Start Time</u>	End Time	Rate Type
1	09:30	10:30	SAVER
1	11:00	12:00	SAVER
1	14:00	15:30	STANDARD
2	10:00	11:30	PREMIUM-B
2	11:30	13:30	PREMIUM-B
2	15:00	16:30	PREMIUM-A

A Rate Type \rightarrow Court függőség igaz, de a meghatározó attribútum (Rate Type) se nem kulcsjelölt se nem szuperkulcs.



A BCNF elérése dekompozícióval



A nem BCNF-ben lévő relációt szét kell bontani:

RATE

Rate Types

<u>Rate Type</u>	Court
SAVER	1
STANDARD	1
PREMIUM-A	2
PREMIUM-B	2

BOOKING

Today's Bookings

<u>Rate Type</u>	<u>Start Time</u>	<u>End Time</u>
SAVER	09:30	10:30
SAVER	11:00	12:00
STANDARD	14:00	15:30
PREMIUM-B	10:00	11:30
PREMIUM-B	11:30	13:30
PREMIUM-A	15:00	16:30

Mindkét reláció BCNF-ben van.

Egyéb normál formák



- **4NF:** Az R relációs séma 4NF-ben van ha valahányszor egy nem triviális (Y nem részhalmaz X -nek, és $X \cup Y$ nem az összes attribútum) $X \twoheadrightarrow Y$ *többértékű függőség* igaz, akkor X egy superkulcs.
- **5NF:** ha 4NF-ben van, és nincs nem triviális *összekapcsolási függősége*.
- **Domain-key NF:** Minden táblára vonatkozó korlátozás *logikai következménye* a táblázat domain korlátozásainak és kulcs korlátozásainak.
- **6NF:** elméleti jelentőségű, nagyon ritka

Normalizáció



PÉLDA

Példa – Egyszerűsített autó biztosítási DB



- Tárolandó adatok:
 - Rendszám
 - Alvázszám
 - A gépjármű típusa (pl. Ford Focus, Mazda 5)
 - Gyártási éve
 - A gépjármű színe
 - Hengerűrtartalom
 - Tulajdonos neve (csak 1 lehet!)
 - Tulajdonos szigszáma
 - Tulajdonos lakcíme
 - Tulajdonos telefonszáma (csak 1-et adhat meg)
 - Biztosító neve
 - Biztosító központi címe
 - Biztosító központi telefonszáma (csak 1)
 - Biztosítás típusa (pl. kötelező, casco)
 - Éves díj
 - Üzletkötő neve

Egyszerűsített autó biztosítás DB – 1NF



- *Tegyünk minden attribútumot bele 1 táblázatba*

AUTÓ_BIZTOSÍTÁS(Rendszám, Alvákszám, Típus, Gyártásiév, Szín,
Hengerűrtartalom, Tulajdonos, Szígszám, Lakcím
Tulajdonos_tel, Biztosító, Biztosító_cím, Biztosító_tel,
Biztosítástípus, Évesdíj, Üzletkötő)

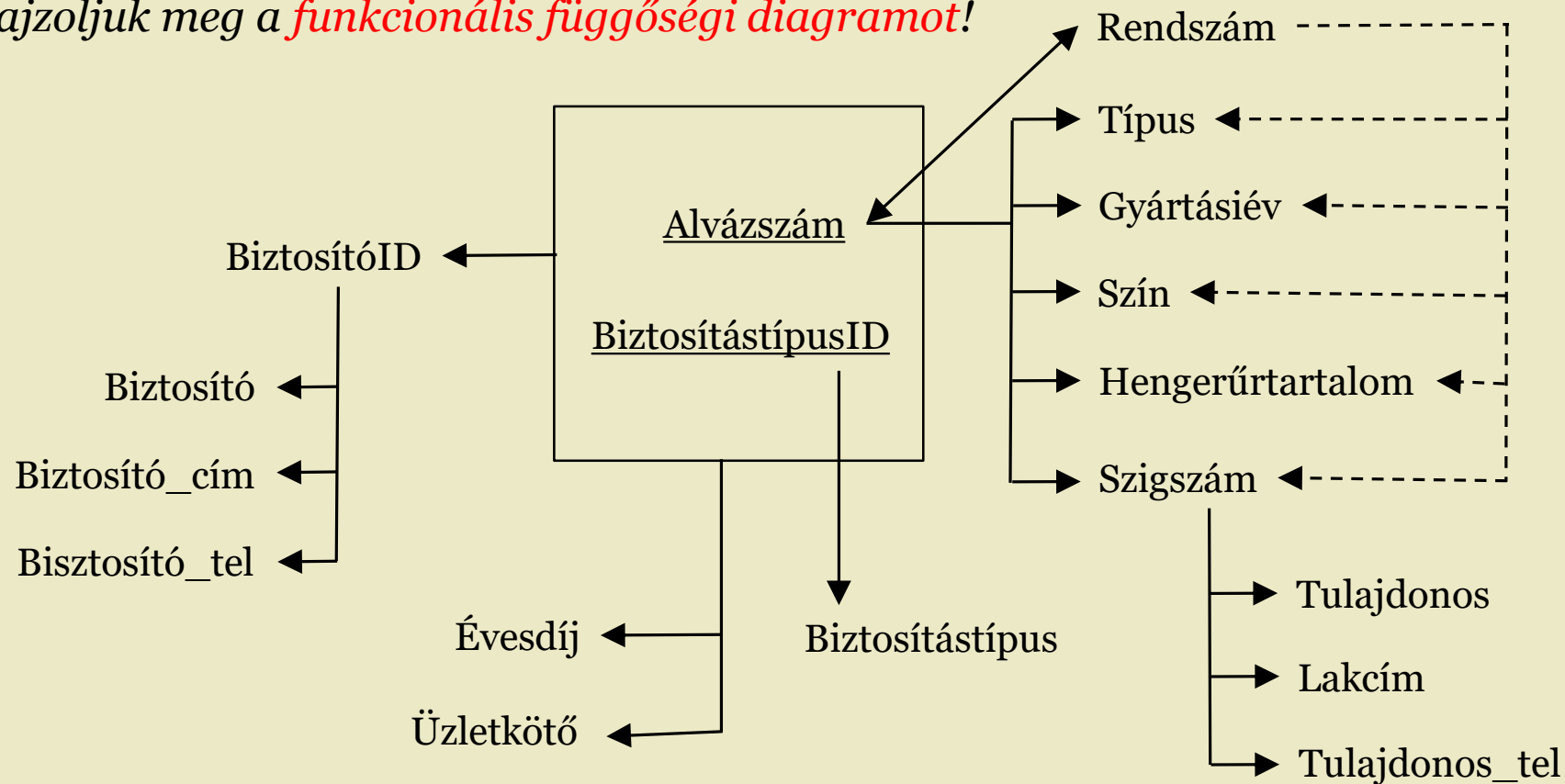
Kódoljuk az egyedeket (BiztosítóID, BiztosítástípusID), és keressük meg az elsődleges kulcsot!

1NF:

AUTÓ_BIZTOSÍTÁS(Rendszám, Alvákszám, Típus, Gyártásiév, Szín,
Hengerűrtartalom, Tulajdonos, Szígszám, Lakcím
Tulajdonos_tel, *BiztosítóID*, Biztosító, Biztosító_cím,
Biztosító_tel, *BiztosítástípusID*, Biztosítástípus, Évesdíj,
Üzletkötő)

Egyszerűsített autó biztosítás (folyt.)

Rajzoljuk meg a *funkcionális függőségi diagramot!*



Egyszerűsített autó biztosítás DB – 2NF



2NF:

AUTÓ (Alvázszám, Rendszám, Típus, Gyártásiév, Szín,
Hengerűrtartalom, Tulajdonos, Szígszám, Lakcím
Tulajdonos_tel)

AUTÓ_BIZTOSÍTÁS(Alvázszám, BiztosítástípusID, BiztosítóID, Biztosító,
Biztosító_cím, Biztosító_tel, Évesdíj, Üzletkötő)

BIZTOSÍTÁSTÍPUS(BiztosítástípusID, Biztosítástípus)

Egyszerűsített autó biztosítás DB – 3NF



3NF:

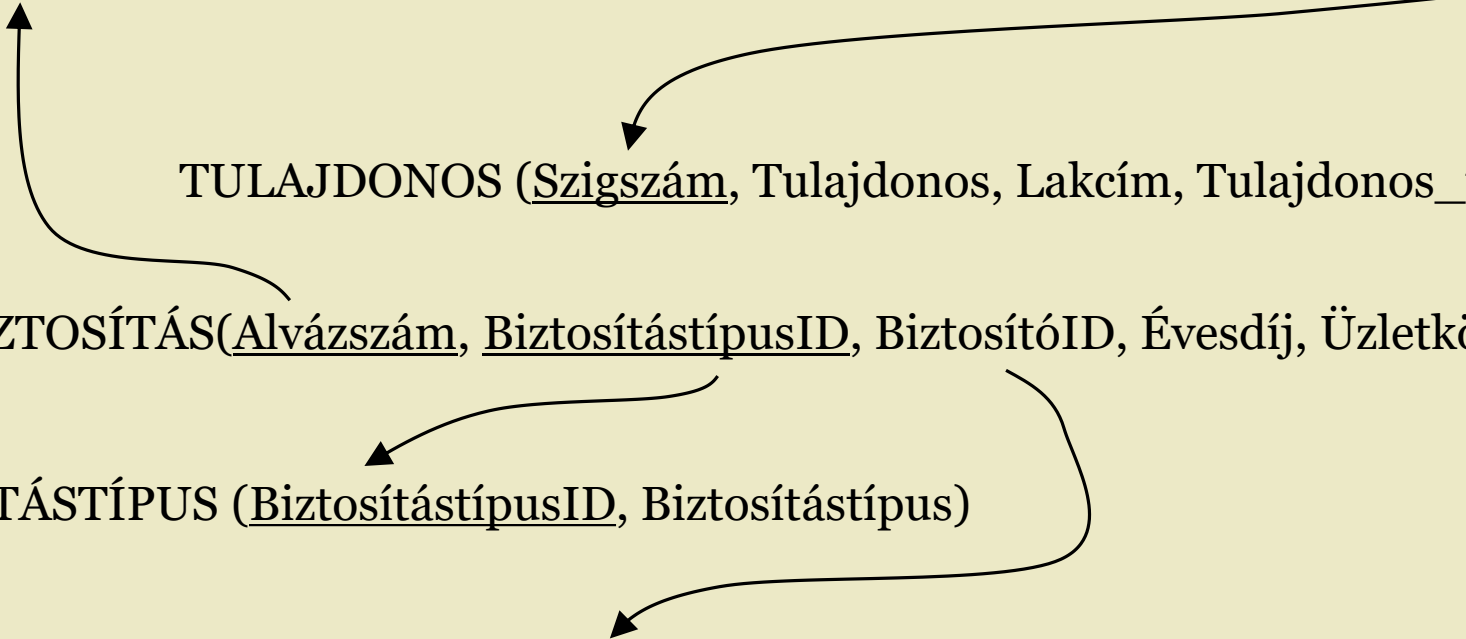
AUTÓ (Alvázszám, Rendszám, Típus, Gyártásiév, Szín, Hengerűrtartalom, Szígszám)

TULAJDONOS (Szígszám, Tulajdonos, Lakcím, Tulajdonos_tel)

AUTÓ_BIZTOSÍTÁS(Alvázszám, BiztosítástípusID, BiztosítóID, Évesdíj, Üzletkötő)

BIZTOSÍTÁSTÍPUS (BiztosítástípusID, Biztosítástípus)

BIZTOSÍTÓ(BiztosítóID, Biztosító, Biztosító_cím, Biztosító_tel)



Egyszerűsített autó biztosítás DB - Kulcsok



- AUTÓ reláció:
 - Elsődleges kulcs: *Alvázsám*
 - Idegen kulcs: *Szigsám* hivatkozik a TULAJDONOS relációra
- TULAJDONOS reláció:
 - Elsődleges kulcs : *Szigsám*
- AUTÓ_BIZTOSÍTÁS reláció:
 - Elsődleges kulcs (összetett): *Alvázsám + BiztosítástípusID*
 - Idegen kulcs 1: *Alvázsám* hivatkozik az AUTÓ relációra
 - Idegen kulcs 2: *BiztosítástípusID* hivatkozik a BIZTOSÍTÁSTÍPUS relációra
 - Idegen kulcs 3: *BiztosítóID* hivatkozik a BIZTOSÍTÓ relációra
- BIZTOSÍTÁSTÍPUS reláció:
 - Elsődleges kulcs : *BiztosítástípusID*
- BIZTOSÍTÓ reláció:
 - Elsődleges kulcs : *BiztosítóID*

Normalizáció



KÖVETELMÉNYEK

Követelmények a normalizációhoz



- A normalizációhoz ismernünk kell:
 - a tárolandó adatok körét, és
 - a rajtuk értelmezett funkcionális függőségeket
 - ✦ *Mennyit?*
 - ✦ *Milyen típusúakat?*

Funkcionális függőség (folyt.)



Tegyük fel, hogy X és Y attribútumhalmazok

- **Triviális funkcionális függőség:**
 - $X \rightarrow Y$ triviális, ha $Y \subseteq X$
- **Nemtriviális funkcionális függőség:**
 - $X \rightarrow Y$ nemtriviális, ha $Y \not\subseteq X$
- **Teljesen nemtriviális funkcionális függőség:**
 - $X \rightarrow Y$ teljesen nemtriviális, ha $Y \cap X = \emptyset$

Attribútumhalmaz lezártja



- **Attribútumhalmaz lezártja** $\{A_1, A_2, \dots, A_n\}^+$
 - Adott $A = \{A_1, A_2, \dots, A_n\}$ attribútumhalmaz
 - Keressük azon B attribútumokat, melyekre igaz: $A \rightarrow B$
 - Az A **attribútumhalmaz lezártja** azon attribútumok halmaza, amelyeket A funkcionálisan meghatároz.

Attribútumhalmaz lezártjának kiszámítása



ALGORITMUS:

- Egy attribútumhalmazból kiindulva az attribútumhalmazt kiterjesztjük azon funkcionális függőségek jobb oldalával, amely funkcionális függőségek bal oldalát már tartalmazza az attribútumhalmaz.
 - Ha a halmaz tovább már nem bővíthető, akkor az eredményhalmaz a **lezárt**.
- 1 Jelölje **X** a lezárt halmazt.
Inicializáljuk X-et a következőképpen: $X = \{A_1, A_2, \dots, A_n\}$.
 - 2 **Ismételjük** minden funkcionális függőségre **S**-ben:
 $B_1 B_2 \dots B_m \rightarrow C$, ahol **B**-k már benne vannak **X**-ben, de **C** nem:
Adjuk **C**-t **X**-hez.
 - 3 **Ismételjük a 2-es lépést** addig amíg nem lehet több attribútumot hozzáadni **X**-hez.
 - 4 Az eredményhalmaz **X** amelyhez már nem tudunk több attribútumot adni tartalmazza az $\{A_1, A_2, \dots, A_n\}$ halmaz lezártját.

Attribútumhalmaz lezártjának kiszámítása – Példa



- Tekintsünk egy relációt a következő attribútumokkal: A, B, C, D, E és F. A reláción a következő funkcionális függőségek igazak:
 - $AB \rightarrow C$,
 - $BC \rightarrow AD$,
 - $D \rightarrow E$,
 - $CF \rightarrow B$.
 - *Mi lesz $\{A,B\}^+$?*
- Iterációk:

– $X = \{A,B\}$	Használjuk: $AB \rightarrow C$
– $X = \{A,B,C\}$	Használjuk : $BC \rightarrow AD$
– $X = \{A,B,C,D\}$	Használjuk : $D \rightarrow E$
– $X = \{A,B,C,D,E\}$	Nincs több lehetséges változás, ezért: $\mathbf{X} = \{A,B\}^+$.
- Az $CF \rightarrow B$ funkcionális függősége nem használható fel, mivel a teljes bal oldalát sohasem tartalmazza \mathbf{X} .

Lezárások és kulcsok



- Adott az R reláció a következő attribútumokkal: $\{A_1, A_2, \dots, A_n\}^+$.
- $\{A_1, A_2, \dots, A_n\}^+ =$ minden attribútum, *akkor és csak akkor*, ha $\{A_1, A_2, \dots, A_n\}$ az R reláció **szuperkulcsa**.
- A_1, A_2, \dots, A_n *a reláció kulcsa?* Ellenőrizzük, hogy:
 - **először**, hogy $\{A_1, A_2, \dots, A_n\}^+$ tartalmazza-e az összes attribútumot,
 - **és nincs** olyan S részhalmaza $\{A_1, A_2, \dots, A_n\}^+$ -nak, hogy S^+ az összes attribútumot adja.

Funkcionális függőségi halmazok lezárása



- Adott a funkcionális függőségek 2 halmaza: S_1 és S_2
- Az S_2 funkcionális függőségek halmaza **levezethető (következik)** az S_1 funkcionális függőségek halmazából, ha minden relációelőfordulás, ami kielégíti az S_1 összes funkcionális függőségét kielégíti egyben az összes S_2 -beli funkcionális függőséget is.
- Példa :
 - S_2 : $\{Szigasz \rightarrow Részleg_neve\}$
 - S_1 : $\{Szigasz \rightarrow RészlegID, RészlegID \rightarrow Részleg_neve\}$
 - S_2 levezethető S_1 -ből
- Hogyan tesztljük?
 - $\{A_1, A_2, \dots, A_n\} \rightarrow \{B_1, B_2, \dots, B_m\}$ levezethető S -ből?
 - Számítsuk ki $\{A_1, A_2, \dots, A_n\}^+$ -t S alapján, és ellenőrizzük, hogy $\{B_1, B_2, \dots, B_m\}$ benne van-e a lezárt halmazban.

Funkcionális függőségek meghatározása a relációs tervezéshez



- Meg kell határozni:

A teljesen nemtriviális függőségeknek egy olyan minimális halmazát, amelyre igaz, hogy belőle minden olyan funkcionális függőség levezethető, ami igaz a reláción.