

Adatbázisrendszerek

6. előadás: Normálformák, normalizálás

Karbantartási anomáliák, 1NF, 2NF, 3NF, BCNF, 4NF, 5NF

2023. március 24.



**DEBRECENI
EGYETEM**





Nem hivatalos tervezési irányelvek relációs adatbázisokhoz

2

6. előadás:
Normálformák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

- Mit jelent a relációs adatbázis-tervezés?
 - Az attribútumok csoportosítását, hogy “jó” relációsémákat alkossanak.
- A relációsémák két szintje:
 - logikai “felhasználói nézet” szint,
 - tárolási (fizikai) “alapreláció” szint.
- A tervezés elsősorban az alaprelációkkal foglalkozik.
- Melyek a “jó” alaprelációk követelményei?
- Először a jó relációs tervezés nem hivatalos irányelveit tárgyaljuk.
- Ezután a **funkcionális függések** és a **normálformák** formális definícióit tekintjük át:
 - 1NF (első normálforma)
 - 2NF (második normálforma)
 - 3NF (harmadik normálforma)
 - BCNF (Boyce–Codd-féle normálforma)
- A függések egyéb típusaival, további normálformákkal az előadás végén foglalkozunk.



1. (nem hivatalos) irányelv

Egy reláció minden egyes rekordja egy egyed-előfordulást vagy kapcsolat-előfordulást reprezentáljon. (Az egyes relációkra és azok attribútumaira külön-külön vonatkozik.)

- Különböző egyedek (DOLGOZÓ-k, OSZTÁLY-ok, PROJEKT-ek) attribútumai nem keverendők egyazon relációban.
- Más egyedekre való hivatkozás csak külső kulcsok használatával történjen.
- Az egyedekre és a kapcsolatokra vonatkozó attribútumokat a lehető legjobban el kell különíteni egymástól.

Tömören

Olyan sémát kell tervezni, ami könnyen magyarázható relációról relációra. Az attribútumok szemantikájának könnyen értelmezhetőnek kell lennie.



Redundáns információk a rekordokban és a karbantartási anomáliák

4

6. előadás: Normálformák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Amikor az információt **redundánsan** tároljuk, az

- tárhelyet pazarol,
- **karbantartási anomáliákat** okoz, amelyek lehetnek
 - beszúrási anomáliák,
 - törlési anomáliák és
 - módosítási anomáliák.

2. (nem hivatalos) irányelv

- Olyan sémát tervezünk, amelyben nem jelennek meg beszúrási, törlési és módosítási anomáliák.
- Ha mégis előfordulnak anomáliák, akkor jegyezzük fel azokat, hogy az alkalmazások számításba vehessék őket.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attríbútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Tekintsük az alábbi relációt:

DOLG_PROJ(Dszsz, Pszám, Dnév, Pnév, Órák)

Példa módosítási anomáliára

Ha megváltoztatjuk a P1 számú projekt nevét “Számlázás”-ról például “Ügyfél-számlakezelés”-re, akkor ezt a módosítást mind a 100 olyan dolgozó esetén végre kell hajtanunk, aki a P1 projekten dolgozik.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Tekintsük az alábbi relációt:

DOLG_PROJ(Dszsz, Pszám, Dnév, Pnév, Órák)

Példa beszűrési anomáliára

Nem tudunk új projektet beszúrni, ha nincs hozzárendelve egyetlen dolgozó sem.

Példa beszűrési anomáliára – megfordítva

Nem tudunk új dolgozót beszúrni, ha nincs hozzárendelve egyetlen projekthez sem.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Tekintsük az alábbi relációt:

DOLG_PROJ(Dszsz, Pszám, Dnév, Pnév, Órák)

Példa törlési anomáliára

Ha törölünk egy projektet, akkor az összes olyan dolgozó is törlődik, aki az adott projekten dolgozik.

Példa törlési anomáliára – a másik oldalról

Ha egy dolgozó egyedüliként dolgozik egy projekten, akkor a dolgozó törlése a szóban forgó projekt törlését is maga után vonja.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

3. (nem hivatalos) irányelv

- A relációkat úgy kell megtervezni, hogy a rekordjaik a lehető legkevesebb NULL értéket tartsanak.
- Azok az attribútumok, amelyek gyakran vesznek fel NULL értéket, külön relációkba tehetők (az elsődleges kulccsal).

A NULL értékek okai:

- Az attribútum nem értelmezhető vagy érvénytelen.
- Az attribútumérték ismeretlen (de létezhet).
- Az érték biztosan létezik, de nem elérhető.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

- A relációs adatbázisok rossz tervezése bizonyos összekapcsolási műveletek esetén hibás eredményhez vezethet.
- A “veszteségmentes összekapcsolás” tulajdonsággal garantáljuk, hogy az összekapcsolási műveletek értelmes eredményt adnak.

4. (nem hivatalos) irányelv

- A relációkat úgy kell megtervezni, hogy kielégítsék a veszteségmentes összekapcsolás feltételét.
- Egy tetszőleges, relációkon végrehajtott természetes összekapcsolás nem állíthat elő álrekordokat.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

A dekompozíciók két fontos tulajdonsággal rendelkeznek:

- A megfelelő összekapcsolás **nemadditív** vagy **veszteségmentes**.
- Megőrzik a funkcionális függéseket.

Jegyezzük meg, hogy

- az (a) tulajdonság különösen fontos, és **nem** áldozható fel,
- a (b) tulajdonság kevésbé szigorú és feláldozható.

Példa álrekordokra:

Dolg_rajta(Szsz, Pszám, Óra, Dnév, Pnév, Phely)

Dolg_helye(Dnév, Phely)

Dolg_proj(Szsz, Pszám, Óra, Pnév, Phely)

Dolg_helye * Dolg_proj $\supset (\neq)$ Dolg_rajta



6. előadás: Normálformák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

A **normalizáció** az a folyamat, amelynek során szétbontjuk a nem kielégítő, “rossz” relációsémákat úgy, hogy az attribútumaikat több kisebb relációsémába helyezzük át.

Definíció

A **normálforma** a relációsémák kulcsai és a bennük fennálló funkcionális függések segítségével megfogalmazott feltétel, amellyel megállapítható, hogy a relációséma egy adott normálformában van-e.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

- 2NF, 3NF, BCNF
a relációsémák **kulcsai** és a bennük fennálló **funkcionális függések** alapján
- 4NF
kulcsok és **többértékű függések** alapján
- 5NF
kulcsok és **join függések** alapján
- További tulajdonságok lehetnek szükségesek a jó relációs tervezés biztosításához (**vesztésmentes összekapcsolás, függésmegőrzés**)



6. előadás: Normálformák

Tervezési
irányelvek

Karbantartási
normálformák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

- A **normalizációt** úgy hajtják végre a gyakorlatban, hogy a kapott tervek magas színvonalúak legyenek, és rendelkezzenek az elvárt tulajdonságokkal.
- A normálformák gyakorlati alkalmazhatósága megkérdőjelezhetővé válik, ha az alapjukat képező megszorítások **nehezen értelmezhetőek** vagy **nehezen ismerhetőek fel**.
- Az adatbázis-tervezőknek **nem szükséges** a lehető legmagasabb normálformáig normalizálniuk (általában csak 3NF-ig, BCNF-ig vagy 4NF-ig).
- A **denormalizáció** az a folyamat, amelynek során magasabb normálformájú relációk összekapcsolását letároljuk alap relációként – alacsonyabb normálformában.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy $R = \{A_1, A_2, \dots, A_n\}$ relációséma **szuperkulcsa** minden olyan $S \subseteq R$ attribútumhalmaz, amelyre igaz, hogy bármely R feletti legális r relációban nincs két olyan t_1 és t_2 rekord, amelyekre $t_1[S] = t_2[S]$ teljesül.

Definíció

A K **kulcs** egy olyan szuperkulcs, amelyből bármely attribútum eltávolítása azt eredményezi, hogy K már nem lesz szuperkulcs többé.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

**Kulcsok és
attribútumok**

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Ha egy relációsémának egynél több kulcsa van, akkor ezeket **kulcsjelölteknek** nevezzük. A kulcsjelöltek közül egy **tetszőlegesen** kiválasztott lesz az **elsődleges kulcs**, a többi másodlagos kulcsoknak nevezzük. Minden relációsémának kell, hogy legyen elsődleges kulcsa.

Definíció

Az R relációséma egy attribútumát R egy **elsődleges attribútumának** nevezzük, ha eleme R **valamely kulcsjelöltjének**. Egy attribútumot **másodlagos (leíró)** attribútumnak hívunk, ha nem elsődleges attribútum, azaz nem eleme egyetlen kulcsjelöltnek sem.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Tiltja

- az összetett attribútumokat,
- a többértékű attribútumokat,
- a **beágyazott relációkat**: az olyan attribútumokat, amelyek értékei a **különálló rekordokban** nem atomiak.

A reláció definíciójának részét képezi.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Hogyan érjük el az első normálformát?

- Távolítsuk el a problémás attribútumot egy másik relációba
- Bővítsük a kulcsot
- Használjunk több atomi attribútumot

Halmazértékű attribútum problémája:

- Új reláció létrehozása az eredeti kulcsával és egy új attribútummal, a kettő együtt alkotja az új reláció kulcsát.
- Az eredeti reláció bővítése annyi új attribútummal amennyi érték a halmazban maximum előfordulhat, a kulcs nem változik. (Sok NULL értéket állíthat elő ezért nem szerencsés.)
- Az eredeti reláció bővítése egy új attribútummal és a kulcs bővítése ezzel az új attribútummal. (Esetenként jelentősen bővítheti a rekordok számát.)

6. előadás:
Normálfor-
mákTervezési
irányelvekKarbantartási
anomáliákNULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálásKulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálformaÖtödik
normálforma

(a)

OSZTÁLY

Onév	<u>Oszám</u>	Ovez_szsz	Ohelyszínek

(b)

OSZTÁLY

Onév	<u>Oszám</u>	Ovez_szsz	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_szsz	Ohelyszínek
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



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy $X \rightarrow Y$ funkcionális függés **teljes funkcionális függés**, ha X -ből bármely A attribútumot eltávolítva a függés a továbbiakban már nem áll fenn, azaz bármely $A \in X$ attribútum esetén $(X - \{A\})$ már **nem** határozza meg funkcionálisan Y -t.

Definíció

Egy $X \rightarrow Y$ funkcionális függés **részleges függés**, ha valamely $A \in X$ attribútum eltávolítható X -ből úgy, hogy a függés továbbra is fennáll, azaz valamely $A \in X$ esetén $(X - \{A\}) \rightarrow Y$.

Definíció

Egy R relációséma **második normálformában** (2NF-ben) van, ha R minden másodlagos (leíró) attribútuma **teljesen funkcionálisan függ** R elsődleges kulcsától.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Az eredeti relációból eltávolítjuk a részelegetesen függő másodlagos (leíró) attribútumot (A) egy másik relációba. Ebben a második relációban szerepelnie kell az eredeti reláció elsődleges kulcsának azon részének (X), amelytől a másodlagos attribútum függ. A második reláció elsődleges kulcsa X , azaz az eredeti reláció elsődleges kulcsának része.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

A harmadik normálforma a **tranzitív függés** fogalmán alapul.

Definíció

Egy R relációséma $X \rightarrow Y$ funkcionális függése **tranzitív függés**, ha létezik egy olyan Z attribútumhalmaz, amely nem kulcsjelölt és nem része R egyetlen kulcsának sem, és fennáll $X \rightarrow Z$ és $Z \rightarrow Y$.

Definíció

Egy R relációséma **harmadik normálformában** (3NF-ben) van, ha 2NF-ben van, **és** nincs R -nek olyan másodlagos (leíró) attribútuma, amely tranzitívan függne az elsődleges kulcstól.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Az eredeti relációból eltávolítjuk a tranzitívan függő másodlagos (leíró) attribútumot egy másik relációba. Ebben a második relációban elsődleges kulcsként kell szerepelnie azoknak az attribútumoknak, amelyekről a másodlagos attribútumok függenek.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy R relációséma **második normálformában** (2NF-ben) van, ha R -nek nincs olyan másodlagos (leíró) attribútuma, amely részlegesen függne R **bármely** kulcsától.

Definíció

Egy R relációséma **harmadik normálformában** (3NF-ben) van, ha valahányszor egy $X \rightarrow A$ **nemtriviális** funkcionális függés fennáll R -en, akkor vagy (a) X egy superkulcsa R -nek, vagy (b) A egy elsődleges attribútuma R -nek.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy R relációséma **Boyce–Codd-féle normálformában** (BCNF-ben) van, ha valahányszor egy $X \rightarrow A$ **nemtriviális** funkcionális függés fennáll R -en, akkor X egy szuperkulcsa R -nek.

Példa

Olyan reláció, amely 3NF-ben van viszont nincs BCNF-ben. Legyen $R(A, B, C)$ reláció AB kulccsal és $FD1 : AB \rightarrow C$, $FD2 : C \rightarrow B$ funkcionális függésekkel. Ekkor R 3NF-ben van, hiszen AB szuperkulcs ($FD1$) és B elsődleges attribútum ($FD2$), viszont nincs BCNF-ben mivel C nem szuperkulcs.



Többértékű függés (multivalued dependency, MVD)

25

6. előadás:
Normálformák

Tervezési irányelvek

Karbantartási anomáliák

NULL értékek a rekordokban

Álrekordok

Normálformák normalizálás

kulcsok és attribútumok

1NF

2NF

3NF

BCNF

Negyedik normálforma

Ötödik normálforma

Informálisan: valahányszor két **független** $1 : N$ számosságú $A : B$ és $A : C$ kapcsolattípust összemixelünk, összemosunk egy $R(A, B, C)$ relációsémába, többértékű függés keletkezhet.

Definíció

Egy R relációsémán megadott $X \twoheadrightarrow Y$ többértékű függés, ahol X és Y R attribútumhalmazai, a következő megszorítást jelenti bármely R fölötti r reláció esetén: Ha van két olyan t_1 és t_2 rekord r -ben, amelyre $t_1[X] = t_2[X]$, akkor léteznie kell két t_3 és t_4 rekordnak is r -ben a következő tulajdonságokkal, ahol Z -t az $(R - (X \cup Y))$ jelölésére használjuk:

- $t_3[X] = t_4[X] = t_1[X] = t_2[X]$.
- $t_3[Y] = t_1[Y]$ és $t_4[Y] = t_2[Y]$.
- $t_3[Z] = t_2[Z]$ és $t_4[Z] = t_1[Z]$.

Valahányszor $X \twoheadrightarrow Y$, azt mondjuk, hogy X **többértékűen meghatározza** Y -t.



Funkcionális és többértékű függések tulajdonságai

26

6. előadás: Normálformák

Tervezési irányelvek

Karbantartási anomáliák

NULL értékek a rekordokban

Álrekordok

Normálformák normalizálás

kulcsok és attribútumok

1NF

2NF

3NF

BCNF

Negyedik normálforma

Ötödik normálforma

- 1 A funkcionális függések reflexivitási szabálya:
Ha $X \supseteq Y$, akkor $X \rightarrow Y$.
- 2 A funkcionális függések augmentívitási szabálya:
 $\{X \rightarrow Y\} \models XZ \rightarrow YZ$.
- 3 A funkcionális függések tranzitívitási szabálya:
 $\{X \rightarrow Y, Y \rightarrow Z\} \models X \rightarrow Z$.
- 4 A többértékű függések komplementer szabálya:
 $\{X \rightarrow Y\} \models \{X \rightarrow (R - (X \cup Y))\}$.
- 5 A többértékű függések augmentívitási szabálya:
Ha $X \rightarrow Y$ és $W \supseteq Z$, akkor $XW \rightarrow YZ$.
- 6 A többértékű függések tranzitívitási szabálya:
 $\{X \rightarrow Y, Y \rightarrow Z\} \models X \rightarrow (Z - Y)$.
- 7 A funkcionális függésből következik a többértékű függés:
 $\{X \rightarrow Y\} \models X \rightarrow Y$.
- 8 A többértékű függésből bizonyos esetekben következik valamiféle funkcionális függés: Ha $X \rightarrow Y$ és létezik olyan W , amelyre (a) $W \cap Y$ üres, (b) $W \rightarrow Z$, és (c) $Y \supseteq Z$, akkor $X \rightarrow Z$.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy $X \rightarrow Y$ többértékű függést **triviális többértékű függésnek** nevezünk, ha vagy (a) Y részhalmaza X -nek, vagy (b) $X \cup Y = R$. Egy olyan többértékű függést, amely sem (a)-t, sem (b)-t nem elégíti ki, **nemtriviális többértékű függésnek** nevezünk.

Megjegyzés

Egy triviális többértékű függés fenn fog állni R **bármilyen r** relációján; azért nevezzük triviálisnak, mert nem határoz meg semmilyen lényeges vagy értelmes (jelentéssel bíró) megszorítást R -re vonatkozóan.



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy R relációséma **negyedik normálformában** (4NF-ben) van, figyelembe véve az F függések halmazát (amely magában foglalja a funkcionális és többértékű függéseket), ha minden F^+ -beli **nemtriviális** $X \twoheadrightarrow Y$ többértékű függés esetén X szuperkulcsa R -nek.



A dekompozíciók veszteségmentes join tulajdonsága

29

6. előadás: Normálformák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Állományok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Az R relációséma egy $D = \{R_1, R_2, \dots, R_m\}$ dekompozíciója **veszteségmentes join tulajdonságú**, figyelembe véve az R -beli F függések halmazát, ha R minden r relációjára, amely kielégíti F -et, fennáll a következő:

$$*(\pi_{R_1}(r), \dots, \pi_{R_m}(r)) = r.$$

A fenti képletben $*$ a természetes összekapcsolást jelöli.

Megjegyzés

Valahányszor felbontunk egy R relációsémát az $R_1 = (X \cup Y)$ és $R_2 = (R - Y)$ relációsémákra egy $X \twoheadrightarrow Y$ többértékű függés alapján, amely fennáll R -en, a dekompozíció veszteségmentes join tulajdonságú lesz.



Relációsémák bináris dekompozíciójának veszteségmentes join tulajdonsága

30

6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attr. bűtumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

A bináris dekompozíciók veszteségmentes join tulajdonságának
tesztelése

Az R_1 és R_2 relációsémák akkor és csak akkor alkotják az R egy
veszteségmentes join dekompozícióját, figyelembe véve a
funkcionális és többértékű függések F halmazát, ha

$$(R_1 \cap R_2) \twoheadrightarrow (R_1 - R_2),$$

vagy – szimmetrikusan – akkor és csak akkor, ha

$$(R_1 \cap R_2) \twoheadrightarrow (R_2 - R_1).$$



6. előadás: Normálformák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy R relációsémán megadott **kapcsolásfüggés** (join dependency, JD) meghatároz egy megszorítást az R bármely r relációjára. A megszorítás azt írja elő, hogy R minden legális r relációjának kell, hogy legyen egy veszteségmentes join dekompozíciója az R_1, R_2, \dots, R_n sémákba; azaz minden ilyen r -re

$$*(\pi_{R_1}(r), \pi_{R_2}(r), \dots, \pi_{R_n}(r)) = r.$$

Az így előírt megszorítást $JD(R_1, R_2, \dots, R_n)$ -nel jelöljük.

Megjegyzés

A többértékű függés olyan speciális esete a kapcsolásfüggésnek, ahol $n = 2$. Azaz $JD(R_1, R_2)$ implikál egy többértékű függést: $(R_1 \cap R_2) \twoheadrightarrow (R_1 - R_2)$ (illetve szimmetrikusan egy $(R_1 \cap R_2) \twoheadrightarrow (R_2 - R_1)$ többértékű függést is).



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

Kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

Definíció

Egy R sémára megadott $JD(R_1, R_2, \dots, R_n)$ kapcsolásfüggés **triviális** kapcsolásfüggés, ha valamely $JD(R_1, R_2, \dots, R_n)$ -beli R_i relációséma egyenlő R -rel.

Definíció

Egy R relációséma **ötödik normálformában** (5NF-ben) van, figyelembe véve funkcionális, többértékű és kapcsolásfüggések egy F halmazát, ha minden F^+ -beli nemtriviális $JD(R_1, R_2, \dots, R_n)$ esetén minden R_i szuperkulcsa R -nek.

6. előadás:
NormálformákTervezési
irányelvekKarbantartási
anomáliákNULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálásKulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálformaÖtödik
normálforma

(a) DOLG

<u>Dnév</u>	<u>Pnév</u>	<u>Hnév</u>
Kovács	X	János
Kovács	Y	Anna
Kovács	X	Anna
Kovács	Y	János

(b) DOLG_PROJEKTEK DOLG_HOZZÁTARTOZÓK

<u>Dnév</u>	<u>Pnév</u>
Kovács	X
Kovács	Y

<u>Dnév</u>	<u>Hnév</u>
Kovács	János
Kovács	Anna

(c) SZÁLLÍTÁS

<u>Sznév</u>	<u>Hozzávaló_név</u>	<u>Proj_név</u>
Kovács	csavar	X projekt
Kovács	anyacsavar	Y projekt
Vágvölgyi	csavar	Y projekt
Lakatos	anyacsavar	Z projekt
Vágvölgyi	szög	X projekt
Vágvölgyi	csavar	X projekt
Kovács	csavar	Y projekt

(d) R_1

<u>Sznév</u>	<u>Hozzávaló_név</u>
Kovács	csavar
Kovács	anyacsavar
Vágvölgyi	csavar
Lakatos	anyacsavar
Vágvölgyi	szög

 R_2

<u>Sznév</u>	<u>Proj_név</u>
Kovács	X projekt
Kovács	Y projekt
Vágvölgyi	Y projekt
Lakatos	Z projekt
Vágvölgyi	X projekt

 R_3

<u>Hozzávaló_név</u>	<u>Proj_név</u>
csavar	X projekt
anyacsavar	Y projekt
csavar	Y projekt
anyacsavar	Z projekt
szög	X projekt

- (a) A DOLG reláció két többértékű függéssel: $Dnév \twoheadrightarrow Pnév$ és $Dnév \twoheadrightarrow Hnév$.

6. előadás:
NormálformákTervezési
irányelvekKarbantartási
anomáliákNULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálásKulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálformaÖtödik
normálforma

(a) DOLG

<u>Dnév</u>	<u>Pnév</u>	<u>Hnév</u>
Kovács	X	János
Kovács	Y	Anna
Kovács	X	Anna
Kovács	Y	János

(b) DOLG_PROJEKTEK DOLG_HOZZÁTARTOZÓK

<u>Dnév</u>	<u>Pnév</u>
Kovács	X
Kovács	Y

<u>Dnév</u>	<u>Hnév</u>
Kovács	János
Kovács	Anna

(c) SZÁLLÍTÁS

<u>Sznév</u>	<u>Hozzávaló_név</u>	<u>Proj_név</u>
Kovács	csavar	X projekt
Kovács	anyacsavar	Y projekt
Vágvölgyi	csavar	Y projekt
Lakatos	anyacsavar	Z projekt
Vágvölgyi	szög	X projekt
Vágvölgyi	csavar	X projekt
Kovács	csavar	Y projekt

(d) R_1

<u>Sznév</u>	<u>Hozzávaló_név</u>
Kovács	csavar
Kovács	anyacsavar
Vágvölgyi	csavar
Lakatos	anyacsavar
Vágvölgyi	szög

 R_2

<u>Sznév</u>	<u>Proj_név</u>
Kovács	X projekt
Kovács	Y projekt
Vágvölgyi	Y projekt
Lakatos	Z projekt
Vágvölgyi	X projekt

 R_3

<u>Hozzávaló_név</u>	<u>Proj_név</u>
csavar	X projekt
anyacsavar	Y projekt
csavar	Y projekt
anyacsavar	Z projekt
szög	X projekt

(b) A DOLG reláció felbontása két 4NF-ben lévő relációra:
DOLG_PROJEKTEK és DOLG_HOZZÁTARTOZÓK.

6. előadás:
NormálformákTervezési
irányelvekKarbantartási
anomáliákNULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálásKulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálformaÖtödik
normálforma

(a) DOLG

<u>Dnév</u>	<u>Pnév</u>	<u>Hnév</u>
Kovács	X	János
Kovács	Y	Anna
Kovács	X	Anna
Kovács	Y	János

(b) DOLG_PROJEKTEK DOLG_HOZZÁTARTOZÓK

<u>Dnév</u>	<u>Pnév</u>
Kovács	X
Kovács	Y

<u>Dnév</u>	<u>Hnév</u>
Kovács	János
Kovács	Anna

(c) SZÁLLÍTÁS

<u>Sznév</u>	<u>Hozzávaló_név</u>	<u>Proj_név</u>
Kovács	csavar	X projekt
Kovács	anyacsavar	Y projekt
Vágvölgyi	csavar	Y projekt
Lakatos	anyacsavar	Z projekt
Vágvölgyi	szög	X projekt
Vágvölgyi	csavar	X projekt
Kovács	csavar	Y projekt

(d) R_1

<u>Sznév</u>	<u>Hozzávaló_név</u>
Kovács	csavar
Kovács	anyacsavar
Vágvölgyi	csavar
Lakatos	anyacsavar
Vágvölgyi	szög

 R_2

<u>Sznév</u>	<u>Proj_név</u>
Kovács	X projekt
Kovács	Y projekt
Vágvölgyi	Y projekt
Lakatos	Z projekt
Vágvölgyi	X projekt

 R_3

<u>Hozzávaló_név</u>	<u>Proj_név</u>
csavar	X projekt
anyacsavar	Y projekt
csavar	Y projekt
anyacsavar	Z projekt
szög	X projekt

(c) A SZÁLLÍTÁS reláció többértékű függések nélkül 4NF-ben van, de nincs 5NF-ben, ha $JD(R_1, R_2, R_3)$ teljesül.

6. előadás:
NormálformákTervezési
irányelvekKarbantartási
anomáliákNULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálásKulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálformaÖtödik
normálforma

(a) DOLG

<u>Dnév</u>	<u>Pnév</u>	<u>Hnév</u>
Kovács	X	János
Kovács	Y	Anna
Kovács	X	Anna
Kovács	Y	János

(b) DOLG_PROJEKTEK DOLG_HOZZÁTARTOZÓK

<u>Dnév</u>	<u>Pnév</u>
Kovács	X
Kovács	Y

<u>Dnév</u>	<u>Hnév</u>
Kovács	János
Kovács	Anna

(c) SZÁLLÍTÁS

<u>Sznév</u>	<u>Hozzávaló_név</u>	<u>Proj_név</u>
Kovács	csavar	X projekt
Kovács	anyacsavar	Y projekt
Vágvölgyi	csavar	Y projekt
Lakatos	anyacsavar	Z projekt
Vágvölgyi	szög	X projekt
Vágvölgyi	csavar	X projekt
Kovács	csavar	Y projekt

(d) R_1

<u>Sznév</u>	<u>Hozzávaló_név</u>
Kovács	csavar
Kovács	anyacsavar
Vágvölgyi	csavar
Lakatos	anyacsavar
Vágvölgyi	szög

 R_2

<u>Sznév</u>	<u>Proj_név</u>
Kovács	X projekt
Kovács	Y projekt
Vágvölgyi	Y projekt
Lakatos	Z projekt
Vágvölgyi	X projekt

 R_3

<u>Hozzávaló_név</u>	<u>Proj_név</u>
csavar	X projekt
anyacsavar	Y projekt
csavar	Y projekt
anyacsavar	Z projekt
szög	X projekt

(d) A SZÁLLÍTÁS reláció felbontása 5NF-ben lévő relációkra:
 R_1 , R_2 , R_3 .



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

- Kezdetben egyetlen (univerzális) relációséma van megadva $R = (A_1, A_2, \dots, A_n)$ (amely az adatbázis összes attribútumát tartalmazza)
- Az univerzális séma minden attribútumának neve egyedi.
- A modellező meghatározza az R -n fennálló funkcionális függések F halmazát.
- Az R univerzális relációsémát relációsémák egy halmazára bontjuk fel $D = \{R_1, R_2, \dots, R_m\}$ (D egy felbontása (dekompozíciója) R -nek)



6. előadás: Normálfor- mák

Tervezési
irányelvek

Karbantartási
anomáliák

NULL
értékek a
rekordokban

Álrekordok

Normálformák
normalizálás

kulcsok és
attribútumok

1NF

2NF

3NF

BCNF

Negyedik
normálforma

Ötödik
normálforma

- Biztosítsuk, hogy R minden attribútuma legalább egy R_i relációs sémában szerepeljen. (Attribútum megőrzés: egy attribútum sem vezethet el.)
- Függőség megőrzés feltétele: minden R -ben megadott $X \rightarrow Y$ funkcionális függés közvetlenül szerepeljen az R_i relációs sémák (ahol R_i szerepel a D felbontásban) egyikében vagy valamely R_j -ben megjelenő függésből levezethető legyen.