

Relációs adatbázis sématervezés I.

EK modell leképezése relációkra. Azonban problémák léphetnek fel, akár a már tárgyalt megszorítások, akár a relációs modell megszorításai miatt.

Fontos, hogy az attribútumokat jól csoportosítsuk:

- logikai csoportosítás
- fizikai, gépi tárolás

Anomáliák:

- Update anomaly Ha módosul egy adat, akkor mindenhol át kell írni.
pl.: Egyetem(név, lakcím, felvett tárgy, jegy)
Ha a fenti táblában változik a hallgató lakcíme, akkor mindenhol át kell írni, ahol szerepel. (ezt kivédjük normálformára hozással)
- Insert anomaly Bizonyos esetekben előfordul, hogy egy adatot nem tudok önmagában bevinni az adatbázisunkba.
pl.: egy tanárt csak akkor tudok bevinni az adatbázisba, ha már van egy tanszék is bent, ahol dolgozhat. Fordítva nem lehetséges.
- Delete anomaly Ha kitöröljük a tanszéket, akkor az ott dolgozókkal is kezdeni kell valamit. Egyes DBMS rendszereknél bizonyos megszorítások miatt csak akkor tudjuk törölni a tanszék táblát, ha az ott dolgozókat is töröltük.

Mi az a NULL érték? Mikor használhatjuk? Mi a jelentése?

- ha létezik az adat, de nem ismerjük
- ha nincs, ha nem elérhető
- ha nem vagyunk az ismeretének a birtokában

Többféle módon tervezhetünk adatbázist:

- top-down
- normalizálás segítségével

Képzeljük el, hogy egy hatalmas adatbázisba minden adatot egy táblában tárolunk. Tehát egy R reláció az összes adatot tartalmazza. Megfelelő módszerekkel több táblára, több logikailag összetartozó attribútum-halmazra

bonthatjuk

Funkcionális függőség:

Definíció: Az $R(A_1, A_2, \dots, A_n)$ sémájú reláció $B \in \{A_1, A_2, \dots, A_n\}$ attribútuma **funkcionálisan függ** az $A_1, A_2, \dots, A_k$ attribútumoktól, ha mindazon esetekben, amikor R valamely sorai megegyeznek az $A_1, A_2, \dots, A_n$ attribútumokon, akkor ezek a sorok megegyeznek a B attribútumon is.

Jelölés: $A_1, A_2, \dots, A_n \rightarrow B$

	X			Y	

1NF, 2NF, 3NF normálformára alakítás

A normalizálás célja: az egyedek tulajdonságaiból kiindulva helyes adatmodell kialakítása

Normálformák:

- 0. normálforma megköveteli, hogy minden egyednek legyen azonosítója
- 1. normálforma 0.NF, és a reláció minden sorában pontosan egy elemi attribútumérték áll. A nem elemi **attribútumot** új attribútumok beiktatásával lehet megszüntetni.

Példa:
A Hallgató(T-szám, Név, Cím, Tantárgy, Jegy) sémájú reláció nem 1NF-ben van, mert a Cím attribútum nem elemi: a város, utca, házszám tulajdonságokból áll. A Tantárgy attribútum pedig több értéket vesz fel.
- 2. normálforma egyben 1.NF, és minden attribútumnak funkcionálisan függeni kell a kulcstól. Ha összetett kulcsunk van, akkor egy-egy tulajdonság részlegesen függhet egy kulcsrészlettől.
- 3. normálforma egyben 2.NF, minden leíró tulajdonság csakis a az azonosítótól függ! (a leíró tulajdonságok között nem lehet funkcionális függés, tranzitív függés kiküszöbölése)

Most vegyünk egy példát normálformára hozásra!

Legyen egy relációs és sémája a következő:

Rendelés(rendelészám, dátum, vevőnév, vevőkód, vevőcím, számlaszám, Áru(áru_azonosító, áru_megnevezés, egységár, mennyiségi_egység, mennyiség), határidő)

Láthatjuk hogy nincs 1NF-ben, mert az Áru beágyazott reláció. Tehát:

Rendelés(rendelészám, dátum, vevőnév, vevőkód, vevőcím, számlaszám, határidő)
Áru(áru_azonosító, áru_megnevezés, egységár, mennyiségi_egység, mennyiség)

Láthatjuk hogy nincs 2NF-ben, mert pl. az áru_azonosító nem határozza meg a mennyiséget.

Áru(áru_azonosító, áru_megnevezés, egységár, mennyiségi_egység)

Tétel(rendelészám, áru_azonosító, mennyiség)

Rendelés(rendelészám, dátum, vevőnév, vevőkód, vevőcím, számlaszám, határidő)

Láthatjuk, hogy nincs 3NF-ben, mert a rendelészám meghatározza a vevőkódot, a vevőkód tranzitív függés által a vevőnevet.

Áru(áru_azonosító, áru_megnevezés, egységár, mennyiségi_egység)

Tétel(rendelészám, áru_azonosító, mennyiség)

Rendelés(rendelészám, dátum, vevőkód, határidő)

Vevő(vevőkód, vevőnév, vevőcím, számlaszám)

Boyce-Codd normálforma (BCNF)

Definíció: Az R reláció BCNF-ben van, ha minden $\alpha \rightarrow \beta$ funkcionális függőségre α szuperkulcs.

Példa:

Legyen **R(A,B,C,D,E)** és egy következő függőségi halmaz: **F={A→D, B→E, DE→C}**

Most ellenőrizzük, hogy a reláció BCNF-ben van-e! AB lenne a kulcs, emiatt nem BCNF. Nincs, még 2.NF-ben sem! (**A→D, B→E**)

Tehát bontsuk fel a következőképpen: R1:=AD ; R2:=ABCE.

R1 az **A→D** miatt BCNF-ben van, de R2 nincs. R2 kulcsa ekkor ABC, így sem BCNF. Ekkor bontsuk fel R2-t: R21:=BE; R22:=ABC. Így már BCNF.