

Adatmodellezés, alapfogalmak

Vassányi István

Alapok

- A helyes modell az információs rendszer későbbi használhatóságának az alapja, olyan mint a jó munkaruha: véd, de nem akadályozza a munkát
- Objektum-orientált modellezés alapfogalmai
 - Objektum-típus (osztály, entitás), ennek vannak attribútumai, kapcsolatai, azonosítója
 - Egyedek (példányok)
- Példa: iroda-dolgozó: egy-sok, dolgozó-kocsi: egy-egy, dolgozó-projekt: sok-sok, projekt-kutatási projekt: egy-egy (specializáció)
- A kétféle egy-egy kapcsolat közti különbség
- Sok-sok kapcsolat esetén a kapcsolat saját attribútumai. 2-nél több entitást is össze lehet kapcsolni

Modelltípusok

- Adatbázis-modellezési szintek:
 - Fogalmi: entitás-kapcsolati modell
 - Logikai: technológia-függő
 - Fizikai: konkrét alkalmazástól és SW/HW környezettől függ
- A tipikus feladat a fogalmi modellből a logikai ab modell előállítás
- Adatbázis-technológiák: relációs, hierarchikus, hálózatos, dokumentumtár
- A relációs technológia a legelterjedtebb üzleti folyamatok támogatására

A relációs technológia

- Elemei: táblák, mezők, rekordok, cellák
- Alapelvek
 - egy cellában csak egy érték lehet
 - nincs kapcsolat nélküli tábla
 - a kulcs a táblában egyedi, lehet több mezős is
- A fogalmi modell megvalósítása relációs technológiával: külső kulcs és kapcsolótábla fogalma
- A külső kulcsot a hivatkozó táblába kell elhelyezni
- A kapcsolótáblának általában összetett kulcsa van, saját kulcsot csak akkor vezetünk be, ha máshonnan hivatkozni akarunk a kapcsolatra. Ebben az esetben is biztosítani kell a kulcsjelölt külső kulcsok kombinált egyediségét unique constraint létrehozásával

Logikai adatbázis-modell készítése

- Módszerek:
 - Intuitíven a feladat szöveges specifikációjából és az adatleltárból
 - A funkcionális függések feltárása után normalizálással az adatleltárból
- Funkcionális függés: az egyik attribútum (mező) minden esetben, egyértelműen meghatározza egy másik értékét. A kulcstól minden mező függ.
- Tranzitív függés: az attribútum nem csak kulcstól függ funkcionálisan, hanem egy kulcstól függő másik attribútumtól is. Ezt el kell kerülni, mert redundanciát okoz és inkonzisztenciát okozhat.
- A normalizálás alaplépése:
 - A tranzitíven függő mezőket új táblába tesszük, melynek a kulcsa az a nem kulcs mező, melytől függenek
 - Az eredeti táblában ez a mező bent marad, mint külső kulcs

A normalizálás szintjei

- 1NF: minden nem-kulcs mező funkcionálisan függ az egyedtípus azonosítójától (a kulcstól)
- 2NF: több mezőből álló kulcsok esetén minden nem-kulcs mező teljes (nem csak részleges) funkcionális függéssel függ a kulcstól
- 3NF: minden nem-kulcs mező teljes (nem csak részleges) funkcionális függéssel függ a kulcstól, és csak attól függ (nincs tranzitív függés)
- A cél a 3NF előállítása
- Esettanulmány:
<http://vassanyi.ginf.hu/ab/tervezes/Video.htm>

OLTP adatbázisok szerkezete

- On-Line Transaction Processing, cél az üzleti folyamat támogatása
- A sok-sok kapcsolat, kapcsolótábla általában eseményhez/tranzakcióhoz kötődik
- Egy tranzakció esetén hópehely, több tranzakció esetén hógolyó a jellemző szerkezet
- A kapcsolótábla a tranzakcionális középpont, a széleken a törzsállományok
- A tranzakcionális középpont és a törzsállományok jellemző különbségei: méret, változási gyakoriság, hivatkozások
- Tanulság: Egy ismeretlen adatbázison pusztán a külső kulcsok követésével meg lehet határozni a lényeges tranzakciókat

Relációs adatbázisok kulcsai és hivatkozásai

- a NULL érték fogalma, NULL-aritmetika: a NULL-ról egy állítás sem igaz, az IS NULL-t kivéve
- A külső kulcs kényszer típusai ON DELETE/ON UPDATE:
 - no action,
 - cascade: ehelyett javasolt a programozott törlés vagy a logikai törlés
 - set null,
 - set default
- A kulcsválasztás elvei
 - nem fogadunk el külső cég által adott azonosítót kulcsnak
 - kulcsjelölteket (candidate key) unique kényszerrel védünk
 - kulcsot nem módosítunk
- A nem összetett kulcs típusa általában: identity (serial)

Adatmodell és GUI kapcsolata

- nem módosítható vagy rejtett mező: kulcs
- editálható szövegdozok: általában kulcstól függő attribútumok
- legördülő listák: törzsállomány-hivatkozások
- jelölőnégyzet plusz szövegdozozal: specializáció
- lenyíló táblázatok vagy listák: egy-sok, sok-sok kapcsolatok

Mező-orientált vagy Rekord-orientált modell

- A rekord-orientált modell az attribútumok halmazát nem rögzíti
- A record-orientált modell előnyei a mező-orientálthoz képest:
 - Flexibilitás, a rendszerben beálló változások követése az adatmodell változtatása nélkül
 - Nincs felesleges adattárolás, NULL értékű mező
- Hátránya:
 - Lényegesen bonyolultabb programozhatóság
 - Különböző adattípusok egy mezőbe kerülhetnek
- A két adatmodell közötti konverzió az adattisztítás, adatföderáció tipikus feladata