Adatbázisrendszerek
5. előadás: A funkcionális függés és jellemzői
Funkcionális függés, Armstrong-axiómák

2024. április 5.

DEBRECENI EGYETEM
INFORMATIKAI KAR

Relációs adatbázis-tervezés

- Olyan módszerek és formális mutatók szükségesek, amelyek segítenek annak eldöntésében, hogy az attribútumok egyik csoportja miért lesz jobb, mint a másik?
- Milyen a jó relációs séma?
- Feltételezzük, hogy a relációt alkotó attribútumoknak jelentésük van.
- A jelentés (szemantika) mondja meg egy relációbeli rekord megfelelő attribútum értékének a jelentését, interpretációját.
- Szemantikus réteg: Hogyan viszonyul az egyik attribútum a másikhoz?
- A funkcionális függés a következő lépés a koncepcionális, azaz nem ad-hoc adatbázistervezés felé.

Figyelmes adatbázis tervezésnél minden szemantikát figyelembe veszünk és az eredményül kapott adatbázis tervnek világos jelentése van.

Univerzális relációséma

A funkcionális függés egy olyan megszorítás, amely az adatbázis két attribútumhalmaza között áll fenn. Tegyük fel, hogy a relációs adatbázissémánknak n attribútuma van: $A_1, A_2, \ldots, A_n$; és gondoljunk az egész adatbázisunkra úgy, hogy azt egyetlen univerzális

$$R = \{ A_1, A_2, \ldots, A_n \}$$

relációsémával írjuk le.

Definíció

Az R két attribútumhalmaza, X és Y között, $X \rightarrow Y$-nal jelölt funkcionális függés előír egy megszorítást azokra a lehetséges rekordokra, amelyek egy R fölötti r relációt alkothatnak. A megszorítás az, hogy bármely két, r-beli t_1 és t_2 rekord esetén, amelyekre $t_1[X] = t_2[X]$ teljesül, teljesülnie kell $t_1[Y] = t_2[Y]$-nak is.

Más szavakkal: egy R relációsémában X akkor és csak akkor határozza meg funkcionálisan Y-t, ha valahányszor $r(R)$ két rekordja megegyezik az X értékeken, szükségszerűen megegyezik az Y értékeken is.

Megjegyzés

- Abból, hogy egy R-re előírt megszorítás szerint bármely $r(R)$ relációpéldányban nem szerepelhet több, mint egy rekord egy adott X értékkel – azaz X egy szuperkulcsa R-nek –, következik $X \to Y$ az R attribútumainak bármely Y részhalmazára (mivel a kulcsmegszorításból következik, hogy egyetlen legális $r(R)$ állapotban sem lehet két olyan rekord, amelyeknek azonosak lennének az X értékeik).
- Ha $X \to Y$ teljesül R-ben, még semmit sem tudunk mondani arról, hogy $Y \to X$ is teljesül-e R-ben. Ha mind $X \to Y$, mind $Y \to X$ teljesül R-ben, akkor kölcsönös funkcionális függésről beszélünk. Ha sem $X \to Y$, sem $Y \to X$ nem teljesül, akkor azt mondjuk, hogy X és Y funkcionálisan független attribútumhalmazok.

Jelölés

Funkcionális függések felírásakor a halmazt jelölő nyitó és záró zárójelek, valamint a halmaz elemeit elválasztó vesszők megállapodás szerint elhagyhatók, ha az attribútumokat egybetűs nevekkel azonosítjuk. (Egyelemű halmazok esetén általában egyébként is elhagyjuk a halmazt jelölő zárójeleket.) Így például

$$\{A, B\} \to \{C\} \text{ helyett } AB \to C,$$

míg

$$\{A, B, C\} \to \{D, E\} \text{ helyett } ABC \to DE$$

írható.

Ha X és Y attribútumhalmazokat jelölnek, akkor a funkcionális függések mindkét oldalán alkalmazható az XY egyszerűsítés a két attribútumhalmaz uniójának jelölésére. Így például

$$X \cup Y \to Z \text{ helyett } XY \to Z$$

írható.

- A relációs adatbázisok egyik legfontosabb fogalma.
- A szemantikának egy tulajdonsága, azaz kizárólag az attribútumok jelentésétől (interpretációjától), amely egy külső dolog, függ.
- Ha a szemantika azt mondja, hogy az attribútumok két halmaza között funkcionális függés van, akkor ezt a függést megszorításként kell specifikálni. Ki kell mondani, le kell írni az adatbázis-kezelő DDL nyelvén.
- A funkcionális függésnek eleget tevő relációsémákat legális kiterjesztésnek nevezzük. (Valójában csak ilyen lehet.) A reláció állapotok szintén csak legálisak lehetnek.
- A funkcionális függés néha automatikusan teljesül, pl. $\{$Állam, Jogosítvány-azonosító$\} \rightarrow$ Személyi szám (USA).

A funkcionális függések tulajdonságai

1. a reflexivitás szabálya: Ha $X \supseteq Y$, akkor $X \to Y$.
2. az augmentivitás szabálya: $\{X \to Y\} \models XZ \to YZ$.
3. a tranzitivitás szabálya: $\{X \to Y,\ Y \to Z\} \models X \to Z$.
4. a dekompozíció szabálya: $\{X \to YZ\} \models X \to Y$.
5. az additivitás szabálya: $\{X \to Y,\ X \to Z\} \models X \to YZ$.
6. a pszeudotranzitivitás szabálya:

$$\{X \to Y,\ WY \to Z\} \models WX \to Z.$$

Definíció

Egy $X \to Y$ funkcionális függés triviális, ha $X \supseteq Y$, egyébként nemtriviális.

Megjegyzés

$X \to A$ és $X \to B$ az additivitás szabálya miatt implikálja $X \to AB$-t, azonban $XY \to A$ nem implikálja szükségképpen sem $X \to A$-t, sem $Y \to A$-t. Viszont $X \to A$ és $Y \to B$-ből következik, hogy $XY \to AB$ az additivitás és tranzitivitás alapján.

- A reflexivitás szabálya szerint egy attribútumhalmaz mindig meghatározza önmagát, vagy saját maga bármelyik részhalmazát.
- Az augmentivitás szabálya szerint egy funkcionális függés mindkét oldalának ugyanazzal az attribútumhalmazzal történő bővítése újabb érvényes funkcionális függést eredményez.
- A tranzitivitás szabálya szerint a funkcionális függések tranzitívak.

- A dekompozíció szabálya azt mondja, hogy egy funkcionális függés jobb oldaláról eltávolíthatunk attribútumokat.
- Az additivitás szabálya szerint funkcionális függések egy

  $$\{\, X \to A_1,\; X \to A_2,\; \ldots,\; X \to A_n \,\}$$

  halmazát összevonhatjuk egyetlen

  $$X \to \{\, A_1, A_2, \ldots, A_n \,\}$$

  funkcionális függéssé.

A reflexivitás bizonyítása

Tegyük fel, hogy $X \supseteq Y$, és hogy léteznek t_1 és t_2 rekordok R valamely r relációjában úgy, hogy $t_1[X] = t_2[X]$. Ekkor $t_1[Y] = t_2[Y]$, mivel $X \supseteq Y$; ezért $X \to Y$-nak teljesülnie kell r-ben.

5. előadás: Funkcionális függés

Adatbázis-tervezés
Funkcionális függés
Tulajdonságok
Tulajdonságok bizonyítása
Armstrong-axiómák
Attribútum-halmaz lezártja
Funkcionális függések ekvivalenciája
Funkcionális függések minimális halmaza

Az augmentivitás bizonyítása (indirekt módon)

Tegyük fel, hogy $X \to Y$ fennáll R egy r relációjában, de $XZ \to YZ$ nem áll fenn. Ekkor léteznie kell t_1 és t_2 rekordoknak úgy, hogy

1. $t_1[X] = t_2[X]$,
2. $t_1[Y] = t_2[Y]$,
3. $t_1[XZ] = t_2[XZ]$ és
4. $t_1[YZ] \neq t_2[YZ]$.

Ez nem lehetséges, mert (3)-ból kapjuk, hogy

5. $t_1[Z] = t_2[Z]$,

míg (2)-ből és (5)-ből kapjuk, hogy

6. $t_1[YZ] = t_2[YZ]$,

ami ellentmond (4)-nek.

A tranzitivitás bizonyítása

Tegyük fel, hogy

1. $X \to Y$ és
2. $Y \to Z$

fennáll egy r relációban. Ekkor tetszőleges t_1 és t_2 r-beli rekordokra, melyekre igaz, hogy $t_1[X] = t_2[X]$, (1) miatt kapjuk, hogy

3. $t_1[Y] = t_2[Y]$;

így (3)-ból és a (2)-es feltevésünkből azt is kapnunk kell, hogy

4. $t_1[Z] = t_2[Z]$;

ezért $X \to Z$-nek fenn kell állnia r-ben.

A dekompozíció bizonyítása

1. $X \rightarrow YZ$ adott.
2. $YZ \rightarrow Y$, felhasználva a reflexivitás szabályát, és tudva, hogy $YZ \supseteq Y$.
3. $X \rightarrow Y$, alkalmazva a tranzitivitás szabályát (1)-re és (2)-re.

Az additivitás bizonyítása

1. $X \to Y$ adott.
2. $X \to Z$ adott.
3. $X \to XY$, alkalmazva az augmentivitás szabályát (1)-re, azt X-szel bővítve; megjegyezve, hogy $XX = X$.
4. $XY \to YZ$, alkalmazva az augmentivitás szabályát (2)-re, azt Y-nal bővítve.
5. $X \to YZ$, alkalmazva a tranzitivitás szabályát (3)-ra és (4)-re.

A pszeudotranzitivitás bizonyítása

1. $X \rightarrow Y$ adott.
2. $WY \rightarrow Z$ adott.
3. $WX \rightarrow WY$, alkalmazva az augmentivitás szabályát (1)-re, azt W-vel bővítve.
4. $WX \rightarrow Z$, alkalmazva a tranzitivitás szabályát (3)-ra és (2)-re.

Definíció

A reflexivitás, az augmentivitás és a tranzitivitás szabályait együtt Armstrong-axiómáknak nevezzük.

William Ward Armstrong 1974-ben bizonyította be, hogy a reflexivitás, az augmentivitás és a tranzitivitás szabálya együtt helyes és teljes.

Helyesség alatt azt értjük, hogy ha adott egy R relációsémán fennálló funkcionális függéseknek egy F halmaza, akkor bármilyen függés, amely levezethető F-ből a három szabály segítségével, fenn fog állni R minden olyan r relációjában, amely kielégíti az F-beli függéseket.

Teljesség alatt azt értjük, hogy a három szabályt mindaddig ismételten alkalmazva, míg már nem kapunk újabb függéseket, előállítható az F-ből levezethető összes lehetséges függés teljes halmaza. Más szavakkal, F-ből kiindulva kizárólag a három szabály alkalmazásával meghatározható az F^+ függések halmaza, amit F lezártjának hívunk.

Az adatbázis-tervező először megadja azon funkcionális függések F halmazát, amelyek könnyen meghatározhatók R attribútumainak a szemantikájából. Ezután az Armstrong-axiómák segítségével további funkcionális függéseket vezet le, amelyek szintén fennállnak R-en. Szisztematikusan ezeket a funkcionális függéseket úgy lehet meghatározni, hogy először meghatározzuk azon X attribútumhalmazokat, amelyek megjelennek valamely F-beli funkcionális függés bal oldalán, és azután meghatározzuk az összes olyan attribútumot, amelyek függnek X-től.

Definíció

Minden egyes X attribútumhalmazra meghatározzuk az attribútumoknak egy olyan X_F halmazát, amelyet X funkcionálisan meghatároz F alapján; X_F-et X F alatti lezártjának nevezzük.

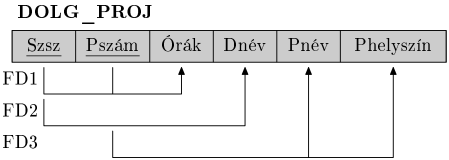
Algoritmus X_F meghatározására

$X_F := X$
repeat
  old$X_F := X_F$;
  for minden F-beli $Y \to Z$ funkcionális függésre **do**
    if $X_F \supseteq Y$ **then** $X_F := X_F \cup Z$;
until $(X_F = \text{old}X_F)$;

Megjegyzés

A fenti algoritmus négyzetes idejű. Létezik egy vele ekvivalens lineáris idejű algoritmus is, amely azonban ennél jóval összetettebb.

Példa X_F meghatározására



Példa

Legyen

$$F = \{\, \{\, \text{Szsz}, \text{Pszám} \,\} \to \text{Órák}, \quad \text{Szsz} \to \text{Dnév}, \quad \text{PSzám} \to \{\, \text{Pnév}, \text{Phelyszín} \,\} \}!$$

Ekkor

$$\{\, \text{Szsz} \,\}_F = \{\, \text{Szsz}, \text{Dnév} \,\},$$
$$\{\, \text{PSzám} \,\}_F = \{\, \text{Pszám}, \text{Pnév}, \text{Phelyszín} \,\},$$
$$\{\, \text{Szsz}, \text{Pszám} \,\}_F = \{\, \text{Szsz}, \text{Pszám}, \text{Dnév}, \text{Pnév}, \text{Phelyszín}, \text{Órák} \,\}.$$

Definíció

Azt mondjuk, hogy a funkcionális függések F halmaza lefedi a funkcionális függések egy másik, E halmazát, ha minden E-beli funkcionális függés benne van F^+-ban; azaz ha minden E-beli függés levezethető F-ből.

Definíció

A funkcionális függések E és F halmaza ekvivalens egymással, ha $E^+ = F^+$. Így az ekvivalencia azt jelenti, hogy minden E-beli funkcionális függés levezethető F-ből, és minden F-beli funkcionális függés levezethető E-ből; azaz E ekvivalens F-fel, ha E lefedi F-et és F lefedi E-t.

Definíció

Funkcionális függések egy F halmazát minimálisnak nevezzük, ha

- Minden $X \to Y$ funkcionális függésre F-ben Y egyszerű, azaz egy attribútumból áll.
- Nem hagyhatunk el egyetlen funkcionális függést sem F-ből úgy, hogy F-fel ekvivalens halmazt kapjunk.
- Nem helyettesíthetünk egyetlen egy $X \to A$ funkcionális függést F-ben egy $Y \to A$ funkcionális függéssel, ahol $Y \subset X$, úgy, hogy F-fel ekvivalens halmazt kapjunk.

Definíció

Funkcionális függések egy E halmazának minimális lefedése alatt (standard kanonikus alakban lévő redundancia mentes) funkcionális függések egy olyan halmazát értjük, amely minimális és ekvivalens E-vel.

Minimális halmaz jellemzői

- Funkcionális függések minden halmazának van vele ekvivalens minimális halmaza.
- Több, egymással is ekvivalens minimális halmaz létezhet.
- A minimális halmaz (nem egyértelmű) standard kanonikus alak, redundanciák nélkül.
- Amikor attribútumok egy halmazából relációkat állítunk elő a tervezés folyamán, akkor feltételezzük, hogy funkcionális függések egy minimális halmazából indulunk ki. Pontosabban, először előállítjuk ezt a minimális halmazt a szemantika által adott funkcionális függésekből.

Input: Funkcionális függések egy E halmaza

1. **Legyen** $F := E$.
2. **Helyettesítsünk** minden $X \to \{A_1, \ldots, A_n\}$ funkcionális függést n számú $X \to A_1$, ..., $X \to A_n$ funkcionális függéssel.
3. Minden olyan F-beli $X \to A$ funkcionális függésre és X-beli B attribútumra, amelyre $\{F \setminus \{X \to A\}\} \cup \{(X \setminus \{B\}) \to A\}$ ekvivalens F-fel, **helyettesítsük** $X \to A$-t $(X \setminus \{B\}) \to A$-val F-ben.
4. Minden fennmaradó F-beli $X \to A$ funkcionális függésre ha $\{F \setminus \{X \to A\}\}$ ekvivalens F-fel, akkor **töröljük** $X \to A$-t F-ből.

Legyen adott $E := \{B \to A, D \to A, AB \to D\}$. Határozzuk meg az F minimális lefedést.

1. Minden funkcionális függés kanonikus alakban van, a 2. lépéssel készen vagyunk.
2. Lehet-e a $AB \to D$-t kiváltani $B \to D$ vagy $A \to D$ valamelyikével. Mivel $B \to A$ így az augmentálás szabálya alapján $B \to AB$, ami a tranzitivitással együtt adja $B \to D$. Ezért $AB \to D$ helyettesíthető $B \to D$-vel. Eredményül $F = \{B \to A, D \to A, B \to D\}$-t kapunk, ahol a baloldalon már mindenütt csak egy attributum szerepel, a 3. lépéssel készen vagyunk.
3. Végül redundáns funkcionális függéseket keresünk F-ben. A tranzitivitás alapján látszik, hogy $B \to D$ és $D \to A$-ból következik $B \to A$, így ez elhagyható. Ezzel végeztünk a 4. lépéssel is.

Eredmény: $F = \{D \to A, B \to D\}$