

Adatbáziskezelés

Funkcionális függőségek

Katona Gyula Y.

- Relációs adatbázistervezés nagy előnye: van elméleti alap
- Kérdés(ek):
 - ▶ Mik a jó relációk?
 - ▶ Milyen relációkat érdemes tárolni?
 - ▶ Hogyan alakíthatunk tetszőleges sémákat jókká?
- Cél: El akarunk kerülni kellemetlen jelenségeket, anomáliákat

- *Módosítási anomália:* pl. ha a Termék(Termelő, Cím, Terméknév, Ár) reláció esetén egy termelő címe több sorban is előfordul, változáskor mindenhol át kell írni. Hiba esetén inkonzisztencia.
- *Beszúrási anomália:* Nem tudunk beszúrni adatot, ha az egyik attribútum hiányzik, mert nem ismerjük (és nem lehet NULL).
- *Törlési anomália:* Csak egész sorok törölhetők, így elveszhetnek hasznos adatok. Pl. ha egy termelő épp nem termel semmit, kitöröljük a címét is.

Mikor jó egy relációs séma?

A relációk, tárolás jósága attól függ, hogy milyen megkötések vannak az adatokon.

Megszorítások két osztálya:

- *Értékfüggő*: Pl. $ÁR \geq 0$, $ÉLETKOR \leq 1000$, $NÉV$ karaktorsor, $CÍM \neq NULL$, (típusleírások)
- *Értékfüggetlen*: $TERMÉKNÉV$, $TERMELŐ$ kulcs; \forall $TERMELŐ$ -nek egy címe van, egy $TERMELŐ$ azonos nevű termékéből csak egy áru van

Utóbbiak: az attribútumok mennyire függenek egymástól \implies **funkcionális függőség**

Funkcionális függőségek

Jelölések: $R(A_1, \dots, A_n)$ reláció, X attribútum halmaz $\implies X \subseteq R$
 $X = \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$ helyett $X = A_{i_1} A_{i_2} \dots A_{i_k}$

Definíció

$Y \subseteq R$ **funkcionálisan függ** $X \subseteq R$ -től, (jelölés: $X \rightarrow Y$), ha R bármely két sorára igaz, hogy ha ők megegyeznek X -en, akkor Y -on is megegyeznek.

PI. $X = \text{TERMELŐ, TERMÉKNÉV}; Y = \text{ÁR} \implies X \rightarrow Y$

Függések típusai

- Azok az érdekes összefüggések, amik minden ilyen attribútumokkal rendelkező táblában mindig fenn kell, hogy álljanak: axiómaszerű feltételek, az adatbázis bármely változása esetén is fennállnak \implies *érdemi függés*
- Azok, amik csak véletlenül, csak egy pillanatban állnak fenn \implies *eseti függés* (ezek nem érdekelnek, például lehetséges hogy egy adott pillanatban minden ár csak egyszer szerepel és ekkor úgy tűnik, mintha Ár \rightarrow Termék érvényes függés lenne)

Relációs séma definíciója

- Tehát az érdemi függések megadása modellezési kérdés: a séma megadásakor döntjük el, hogy milyen függéseket akarunk fenntartani mindenáron.
- Ezentúl a relációs sémának része lesz a függőségek halmaza F is $\implies (R, F)$
Vagyis megadjuk, hogy mik a séma attribútumai és mik az érdemi függései.

Funkcionális függőségek, példák

R(TERMELŐ, TERMÉKNÉV, ÁR, CÍM)

TERMELŐ, TERMÉKNÉV → TERMELŐ, TERMÉKNÉV, ÁR, CÍM
TERMELŐ → CÍM

S(NÉV, CÍM, VÁROS, IRÁNYÍTÓSZ, TELEFON)

CÍM, VÁROS → IRÁNYÍTÓSZ
IRÁNYÍTÓSZ → VÁROS

Funkcionális függések

- Egy adott reláció adott állapotából nem következik semmilyen érdemi függés. Viszont látszódhat olyan, hogy mi nem függhet mitől.
- $X \rightarrow Y$ teljesülhet úgy is, hogy az adott relációban nincs is két olyan sor, amik X -en megegyeznek.
- X -nek és Y -nak nem kell diszjunktaknak lenniük

A séma megadása csak a keretet jelenti, beleértve a függéseket is, ha ezt feltöltjük adatokkal, akkor kapunk egy a sémára illeszkedő relációt. A r reláció akkor illeszkedik az (R, F) sémára ha az attribútumai az R -ben adottak és teljesülnek benne az F függések.

Kérdés

Ha adott egy F függéshalmaz és egy reláció, amiben F függései igazak, akkor milyen további függések lesznek még biztosan igazak?

Például: ha $HALLGATÓ, TÁRGY \rightarrow GYAKORLAT$ és $GYAKORLAT \rightarrow GYAKVEZ$, akkor $HALLGATÓ, TÁRGY \rightarrow GYAKVEZ$.

Azaz általánosabban: ha $XY \rightarrow Z$ és $Z \rightarrow W$, akkor attól függetlenül, hogy mi a reláció és mi X, Y, Z, W , igaz lesz, hogy $XY \rightarrow W$.

Definíció

Adott (R, F) relációs séma. Az $X \rightarrow Y$ függés **logikai következménye** (szemantikai következménye) F -nek, ha az $X \rightarrow Y$ minden olyan r relációban teljesül, ahol F függései mind teljesülnek.

Jelölése: $F \models X \rightarrow Y$

Azaz ez a fogalom azt adja meg, hogy mely függéseknek kell szükségszerűen teljesülniük minden olyan sémában/relációban, ahol F függései fennállnak.

Hogyan lehetne ezeket meghatározni, illetve eldönteni, hogy egy függés ilyen-e?

Kérdés

Hogyan lehet eldönteni, hogy egy függés logikai következménye-e egy F függéshalmaznak?

- Felveszünk axiómákat, és azok segítségével próbálunk új függéseket levezetni F -ből. Azt nézzük, hogy mely függések vezethetők le F -ből.
- Persze ehhez az kell, hogy pontosan azokat lehessen levezetni F -ből, amik logikai következményei neki.
- **Levezethetőség jele:** $F \vdash X \rightarrow Y$

Logikai következmény vs. levezethetőség

Mindjárt bevezetünk axiómákat (ezekkel pedig levezethetőséget) és azt lehet belátni, hogy $\models \iff \vdash$.

(Pl. logikában így van.)

- $\models \Rightarrow \vdash$: *Teljességi tétel*, azaz ami igaz az levezethető.
- $\vdash \Rightarrow \models$: *Igazság tétel*, azaz csak igaz dolgok vezethetők le.

Definíció

Egy $X \rightarrow Y$ függőség akkor vezethető le egy adott F függőséghalmazból, ha az alábbi axiómák véges sokszori ismételt alkalmazásával F -ből megkapjuk $X \rightarrow Y$ -t. Jele: $F \vdash X \rightarrow Y$.

- 1 Reflexivitás: Ha $X, Y \subseteq R$ és $Y \subseteq X$, akkor $X \rightarrow Y$.
- 2 Kiegészítési tulajdonság: Ha $X, Y \subseteq R$ és $X \rightarrow Y$, akkor $XW \rightarrow YW$ igaz tetszőleges $W \subseteq R$ -re.
- 3 Transzitivitás: Ha $X, Y, Z \subseteq R$, $X \rightarrow Y$ és $Y \rightarrow Z$, akkor $X \rightarrow Z$.

Igazságtétel bizonyítása ($\vdash \Rightarrow \models$)

- Azt kell belátni, hogy ha egy függés (esetleg több lépésben) levezethető F -ből a három axióma segítségével, akkor ez a függés logikai következménye is F -nek, azaz minden olyan relációban, ahol F minden függése teljesül, ott teljesül a levezetett függés is.
- Ehhez elég azt belátni, hogy külön-külön, az egyes axiómák egyszeri használatakor ez igaz.
- Vagyis mindhárom axiómát meg fogjuk most nézni.

Igazságtétel bizonyítása (2)

- 1 Reflexivitás: Azt kell belátni, hogy minden r relációban, minden $Y \subseteq X \subseteq R$ attribútumhalmaz esetén $X \rightarrow Y$ igaz, azaz ha r bármely két adott sora megegyezik X -en, akkor megegyeznek Y -on is. De mivel $Y \subseteq X$, ezért nyilván megegyeznek Y -on, ha X -en megegyeztek.
- 2 Kiegészítési tulajdonság: Az kell, hogy ha egy R -re illeszkedő r relációban $X \rightarrow Y$ igaz, akkor $XW \rightarrow YW$ is igaz lesz. Vegyünk két sort r -ben, ami megegyezik XW -n. Ekkor ezek megegyeznek X -en és W -n is, külön-külön. Mivel $X \rightarrow Y$, így megegyeznek Y -n is, tehát YW -n is.

Igazságtétel bizonyítása (3)

- 1 **Tranzitivitás:** Az kell, hogy ha egy R -re illeszkedő r relációban $X \rightarrow Y$ és $Y \rightarrow Z$ igaz, akkor $X \rightarrow Z$ is igaz lesz. Vegyünk két sort, ami megegyezik X -en. Mivel $X \rightarrow Y$, megegyeznek Y -n is. De mivel $Y \rightarrow Z$, megegyeznek Z -n is.

Vagyis készen vagyunk, mert mindhárom axiómára igaz, hogy az axióma bal oldalának logikai következménye a jobb oldala, így az axiómák véges sokszori alkalmazása esetén is igaz lesz, hogy a kiindulási függéshalmaz logikai következménye a levezetett függés.

Példa

Ha $R(\text{Város, Utca, Irányítószám})$ és $F = \{VU \rightarrow I, I \rightarrow V\}$, akkor $F \vdash IU \rightarrow VIU$ (és mivel $\vdash \Rightarrow \models$ -t már láttuk, ezért $F \models IU \rightarrow VIU$).

- i) $I \rightarrow V$: ez F -beli
- ii) $IU \rightarrow VU$: kiegészítve U -val
- iii) $IU \rightarrow IVU$: kiegészítve I -vel

Levezethető szabályok

Néhány további szabály, ami levezethető az axiómákból (és az igazságtétel miatt igazak is.) Mivel ezek levezethetők az axiómákból, ezeket is használhatjuk mostantól levezetések során.

[Unió szabály] $\{X \rightarrow Y, X \rightarrow Z\} \vdash X \rightarrow YZ$

- i) $X \rightarrow Y$: ez F -beli
- ii) $XZ \rightarrow YZ$: kiegészítve Z -val
- iii) $X \rightarrow Z$: ez F -beli
- iv) $X \rightarrow XZ$: kiegészítve X -vel
- v) $X \rightarrow YZ$: iv) és ii) + tranzitivitás

[Áltranzitív szabály] $\{X \rightarrow Y, YW \rightarrow Z\} \vdash XW \rightarrow Z$

- i) $X \rightarrow Y$: ez F -beli
- ii) $XW \rightarrow YW$: kiegészítve W -val
- iii) $YW \rightarrow Z$: ez F -beli
- iv) $XW \rightarrow Z$: ii) és iii) + tranzitivitás

Levezethető szabályok

[Felbontási szabály] Tegyük fel, hogy $Z \subseteq Y$, ekkor $\{X \rightarrow Y\} \vdash X \rightarrow Z$

- i) $X \rightarrow Y$: ez F -beli
- ii) $Y \rightarrow Z$: reflexivitás
- iii) $X \rightarrow Z$: i) és ii) + tranzitivitás

Teljességi tétel

Az igazságtétel fordítottja: azaz $F \models X \rightarrow Y \Rightarrow F \vdash X \rightarrow Y$.

Ez is igaz, nem bizonyítjuk.

Definíció

Ha F egy függéshalmaz, akkor a lezártja (jele F^+) az F -ből levezethető összes függés:

$$F^+ = \{X \rightarrow Y \mid F \vdash X \rightarrow Y\}$$

Jó: mert $\models \iff \vdash$ miatt ez éppen az F -ből szükségszerűen következő összes függést adja meg.

Gond: nagyon nagy lehet

Pl. $R(A_1, \dots, A_n, B_1, \dots, B_n)$ és $F = \{A_i \rightarrow B_j \mid 1 \leq i, j \leq n\}$, akkor ez legalább 2^n db függésmert F^+ -ban benne van minden $A_{i_1} \dots A_{i_k} \rightarrow B_{j_1} \dots B_{j_l}$, azaz $(2^n - 1)(2^n - 1) \approx 2^{2n}$ eleme van.

Attribútumhalmaz lezárása

Ezért ehelyett valami mást nézünk, amit könnyebb lesz meghatározni és jól közelíti F^+ -t:

Definíció

Ha X egy attribútum halmaz (R, F) -ben, akkor X **lezártja**

$$X^+(F) = \{A \in R \mid F \vdash X \rightarrow A\},$$

azaz azon attribútumok, amik függnek X -től.

Nyilván igaz, hogy $X \subseteq X^+(F) \subseteq R$

Lemma

(Fontos!!!) $F \vdash X \rightarrow Y \iff Y \subseteq X^+(F)$

\implies : Tegyük fel, hogy $F \vdash X \rightarrow Y$ és legyen $A \in Y$.
 $F \vdash X \rightarrow A$, hiszen vegyük $X \rightarrow Y$ levezetését és alkalmazzuk a felbontási szabályt a végén.

Definíció szerint ekkor $A \in X^+(F)$. Ez minden $A \in Y$ -ra igaz.

\impliedby : Legyen $Y = A_1 \dots A_k \subseteq X^+(F)$.

Így definíció szerint $\forall A_i \in Y$ -ra $F \vdash X \rightarrow A_i$.

Ekkor $X \rightarrow Y$ levezetése: vesszük az A_i -k levezetését és a végén alkalmazzuk az unió szabályt $k - 1$ -szer.

Fontos lemma következménye

- Ha minden X -re ismerjük/ki tudjuk számítani $X^+(F)$ -et, akkor tetszőleges $X \rightarrow Y$ függésről eldönthető, hogy F^+ -beli-e vagy sem,
- mert $X \rightarrow Y \in F^+$ pontosan akkor teljesül (definíció szerint), ha $F \vdash X \rightarrow Y$, de ez meg az előbbi lemma szerint pontosan akkor van, ha $Y \subseteq X^+(F)$
- Mindjárt látjuk, hogy $X^+(F)$ kiszámolására lesz gyors algoritmus.

$X^+(F)$ kiszámítása

Algoritmus:

$$X_0 = X,$$

⋮

$$X_i = \dots,$$

$$X_{i+1} = X_i \cup \{A \in R \mid \text{van olyan } U \rightarrow V \in F, \text{ hogy } U \subseteq X_i \text{ és } A \in V\},$$

⋮

$$X^+(F) = X_{\text{utolsó}}, \text{ (amikor már nem nő)}$$

Ezt persze be kéne látni, de ezt most kihagyjuk :)

Példa

$R(A, B, C, D)$, $F = \{A \rightarrow B, B \rightarrow C, BC \rightarrow D\}$, $A^+(F) = ?$
 $X_0 = \{A\}$, $X_1 = \{A, B\}$, $X_2 = \{A, B, C\}$, $X_3 = \{A, B, C, D\} = X_{\text{utolsó}}$

Definíció

Definíció:

$X \subseteq R$ **szuperkulcsa** az (R, F) sémának, ha $F \vdash X \rightarrow R$. Másképpen, ha $R = X^+(F)$.
 $X \subseteq R$ **kulcsa** az (R, F) sémának, ha szuperkulcs és nincs olyan valódi részhalmaza, ami szuperkulcs.

Példa:

$F = \{\text{TERMELŐ, TERMÉKNÉV} \rightarrow \text{ÁR}; \text{TERMELŐ} \rightarrow \text{CÍM}\}$

$X = \text{TERMELŐ, TERMÉKNÉV}$

$\implies X^+(F) = \text{TERMELŐ, TERMÉKNÉV, ÁR, CÍM}$

$\text{TERMELŐ}^+(F) = \text{TERMELŐ, CÍM}$

$\text{TERMÉKNÉV}^+(F) = \text{TERMÉKNÉV}$

$\implies X$ kulcs