Adatbáziskezelés
Funkcionális függőségek

Csima Judit

BME, VIK,
Számítástudományi és Információelméleti Tanszék

2017. november 3.

Relációs sémák tervezése

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

Anomáliák

- *Módosítási anomália:* pl. ha a Termék(Termelõ, Cím, Terméknév, Ár) reláció esetén egy termelõ címe több sorban is elõfordul, 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 törölhetõk, így elveszhetnek hasznos adatok. Pl. ha egy termelõ 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:

- *Értékfüggõ:* Pl. ÁR$\geq$ 0, ÉLETKOR egész $\leq$ 1000, NÉV karaktersor, CÍM$\neq$NULL, (típusleírások)
- *Értékfüggetlen:* TERMÉKNÉV, TERMELÕ kulcs; $\forall$ TERMELÕ-nek egy címe van, egy TERMELÕ azonos nevû termékébõl csak egy árú van

Utóbbiak: az attribútumok mennyire függenek egymástól

$\Longrightarrow$ **funkcionális függõség**

Funkcionális függõségek

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

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

Pl. $X =$TERMELÕ, TERMÉKNÉV; Y=ÁR $\Longrightarrow X \rightarrow Y$

Függések típusai

- Azok az érdekes összefüggések, amik minden ilyen attribútumokkal rendelkezõ táblában mindig fenn kell, hogy álljanak: axiómaszerû feltételek, az adatbázis bármely változása esetén is fennállnak
  $\Longrightarrow$ *érdemi függés*
- Azok, amik csak véletlenül, csak egy pillanatban állnak fenn
  $\Longrightarrow$ *eseti függés*
  (ezek nem érdekelnek, például lehetséges hogy egy adott pillanatban minden ár csak egyszer szerepel és ekkor úgy tûnik, 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 függõségek halmaza F is $\Longrightarrow (R, F)$
  Vagyis megadjuk, hogy mik a séma attribútumai és mik az érdemi függései.

Funkcionális függõségek, példák

R(TERMELÕ, TERMÉKNÉV, ÁR, CÍM)

TERMELÕ, TERMÉKNÉV $\rightarrow$ TERMELÕ, TERMÉKNÉV, ÁR, CÍM
TERMELÕ $\rightarrow$ CÍM

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

CÍM, VÁROS $\rightarrow$ IRÁNYÍTÓSZ
IRÁNYÍTÓSZ $\rightarrow$ VÁROS

Funkcionális függések

- Egy adott reláció adott állapotából nem következik semmilyen érdemi függés.
  Viszont látszódhat olyan, hogy mi nem függhet mitõl.
- $X \rightarrow Y$ 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 függeseket is, ha ezt feltöltjük adatokkal, akkor kapunk egy a sémára illeszkedõ relációt. A r reláció akkor illeszkedik az (R, F) sémára ha az attribútumai az R-ben adottak és teljesülnek benne az F függések.

Logikai következmény

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 $XY \rightarrow W$.

Logikai következmény

Definíció:

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

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

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

Levezethetőség

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-bõl, amik logikai következményei neki.
- **Levezethetõség jele:** $F \vdash X \to Y$

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 levezethetõ.
- $\vdash \Rightarrow \models$: *Igazság tétel*, azaz csak igaz dolgok vezethetõk le.

Armstrong axiómák

Definíció Egy $X \to Y$ függõség akkor vezethetõ le egy adott F függõséghalmazból, ha az alábbi axiómák véges sokszori ismételt alkalmazásával F-bõl megkapjuk $X \to Y$-t. Jele: $F \vdash X \to Y$.

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

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

- Azt kell belátni, hogy ha egy függés (esetleg több lépésben) levezethetõ F-bõl 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 $X \to Y$ 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 illeszkedõ r relációban $X \to Y$ igaz, akkor $XW \to YW$ is igaz lesz. Vegyünk két sort r-ben, ami megegyezik XW-n. Ekkor ezek megegyeznek X-en és W-n is, külön-külön. Mivel $X \to Y$, így megegyeznek Y-n is, tehát YW-n is.

Igazságtétel bizonyítása (3)

1. Tranzitivitás: Az kell, hogy ha egy R-re illeszkedõ 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 $R(\text{Város}, \text{Utca}, \text{Irányítószám})$ és $F = \{VU \to I, I \to V\}$, akkor
$F \vdash IU \to VIU$
(és mivel $\vdash \Rightarrow \models$-t már láttuk, ezért $F \models IU \to VIU$).

i) $I \to V$: ez F-beli

ii) $IU \to VU$: kiegészítve U-val

iii) $IU \to IVU$: kiegészítve I-vel

Levezethetõ szabályok

Néhány további szabály, ami levezethetõ az axiómákból (és az igazságtétel miatt igazak is.) Mivel ezek levezethetők az axiómákból, ezeket is használhatjuk mostantól levezetések során.
[Unió szabály] $\{X \to Y,\ X \to Z\} \vdash X \to YZ$

i) $X \to Y$: ez F-beli

ii) $XZ \to YZ$: kiegészítve Z-val

iii) $X \to Z$: ez F-beli

iv) $X \to XZ$: kiegészítve X-vel

v) $X \to YZ$: iv) és ii) + tranzitivitás

Levezethetõ szabályok

[Áltranzitiv szabály] $\{X \to Y,\ YW \to Z\} \vdash XW \to Z$

i) $X \to Y$: ez F-beli

ii) $XW \to YW$: kiegészítve W-val

iii) $YW \to Z$: ez F-beli

iv) $XW \to Z$: ii) és iii) + tranzitivitás

Levezethetõ szabályok

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

i) $X \to Y$: ez F-beli

ii) $Y \to Z$: reflexivitás

iii) $X \to Z$: i) és ii) + tranzitivitás

Teljességi tétel

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

Ez is igaz, majd mindjárt belátjuk, csak előbb még kell pár fogalom hozzá.

Függéshalmaz lezárása

Definíció Ha F egy függéshalmaz, akkor a lezártja (jele F^+) az F-bõl levezethetõ összes függés:

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

Jó: mert $\models \iff \vdash$ miatt ez éppen az F-bõl szükségszerûen következõ összes függést adja meg.

Gond: nagyon nagy lehet
Pl. $R(A_1, \ldots, A_n, B_1, \ldots B_n)$ és $F = \{A_i \to B_j \mid 1 \leq i, j \leq n\}$, akkor ez legalább 2^n db függésmert F^+-ban benne van minden $A_{i_1} \ldots A_{i_k} \to B_{j_1} \ldots B_{j_l}$, 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**

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

azaz azon attribútumok, amik függnek X-től.

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

Attribútumhalmaz lezárása

Lemma

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

$\boxed{\Longrightarrow:}$ Tegyük fel, hogy $F \vdash X \to Y$ és legyen $A \in Y$.
$F \vdash X \to A$, hiszen vegyük $X \to Y$ levezetését és alkalmazzuk a felbontási szabályt a végén.
Definíció szerint ekkor $A \in X^+(F)$. Ez minden $A \in Y$-ra igaz.
$\boxed{\Longleftarrow:}$ Legyen $Y = A_1 \dots A_k \subseteq X^+(F)$.
Így definíció szerint $\forall A_i \in Y$-ra $F \vdash X \to A_i$.
Ekkor $X \to Y$ levezetése: vesszük az A_i-k levezetését és a végén alkalmazzuk az uniós szabályt $k-1$-szer.

Fontos lemma következménye

- Ha minden X-re ismerjük/ki tudjuk számítani $X^+(F)$-et, akkor tetszõleges $X \to Y$ függésrõl eldönthetõ, hogy F^+-beli-e vagy sem,
- mert $X \to Y \in F^+$ pontosan akkor teljesül (definíció szerint), ha $F \vdash X \to Y$, de ez meg az elõbbi lemma szerint pontosan akkor van, ha $Y \subseteq X^+(F)$
- Majd látjuk, hogy $X^+(F)$ kiszámolására lesz gyors algoritmus.

Teljességi tétel, azaz $F \models X \to Y \Rightarrow F \vdash X \to Y$.

Tegyük fel indirekt, hogy van olyan $X \to Y$ függés és F függéshalmaz, hogy $X \to Y$ nem vezethetõ le F-bõl ($F \nvdash X \to Y$), noha logikai következménye neki ($F \models X \to Y$).
Ez utóbbi azt jelenti, hogy minden olyan relációban, amiben F függõségei teljesülnek, ha X-en megegyezik két sor, akkor azoknak meg kell egyezniük Y-on is.
Úgy jutunk ellentmondásra, hogy $F \nvdash X \to Y$ felhasználásával konstruálunk egy olyan r relációt, ahol F függõségei teljesülnek, de $X \not\to Y$, ami ellentmond $F \models X \to Y$-nak.

Bizonyítás (folyt.)

r							$X^+(F)$						
							X						
	A_1	…	…				……………				…	…	A_n
t_1	1	1	1	1	1	1	1 ………… 1	1	1	1	1	1	1
t_2	0	0	0	1	1	1	1 ………… 1	1	1	1	0	0	0

Tehát r két soros reláció, $X^+(F)$-en megegyezik a két sor, a többin különbözik.
Meg fogjuk mutatni, hogy r-ben teljesülnek F függései , de r-ben nem igaz $X \to Y$. amivel megkapjuk az ellentmondást.

Bizonyítás: r-ben teljesülnek F függései

$\boxed{\text{r}}$			$X^+(F)$		
			X		
	A_1		....		... ... A_n
t_1	1 1 1	1 1 1	1 1	1 1 1	1 1 1
t_2	0 0 0	1 1 1	1 1	1 1 1	0 0 0

Legyen $U \to V \in F$.
Ha $U \not\subseteq X^+(F) \implies U \to V$ igaz, hiszen nincs olyan két sor, ami U-n megegyezik.
Ha $U \subseteq X^+(F)$, akkor lemma miatt $F \vdash X \to U$.
Tranzitivitás miatt $F \vdash X \to V$. Lemma $\implies V \subseteq X^+(F)$, V-n megegyezik a két sor.

Bizonyítás: r-ben nem igaz $X \to Y$

r							$X^+(F)$						
							X						
	A_1	…	…				....				…	…	A_n
t_1	1	1	1	1	1	1	1 1	1	1	1	1	1	1
t_2	0	0	0	1	1	1	1 1	1	1	1	0	0	0

Mivel $F \not\models X \to Y$, ezért a fontos lemma miatt $Y \not\subseteq X^+(F)$, azaz Y kilóg $X^+(F)$-bõl, abból a részbõl, ahol a két sor egyenlõ.
Vagyis a két sor egyenlõ X-en, de nem egyenlõ Y-on, így $X \to Y$ nem igaz r-ben.

Kulcs

Definíció:
$X \subseteq R$ **szuperkulcsa** az (R, F) sémának, ha $F \vdash X \to R$. Másképpen, ha $R = X^+(F)$.
$X \subseteq R$ **kulcsa** az (R, F) sémának, ha szuperkulcs és nincs olyan valódi részhalmaza, ami szuperkulcs.
Példa:
$F = \{$TERMELÕ, TERMÉKNÉV $\to$ ÁR; TERMELÕ $\to$ CÍM$\}$
$X =$ TERMELÕ, TERMÉKNÉV
$\Longrightarrow X^+(F) =$ TERMELÕ, TERMÉKNÉV, ÁR, CÍM
TERMELÕ$^+(F) =$ TERMELÕ, CÍM
TERMÉKNÉV$^+(F) =$ TERMÉKNÉV
$\Longrightarrow X$ kulcs

$X^+(F)$ kiszámítása

Algoritmus:
$X_0 = X$,
$\vdots$
$X_i = \ldots,$
$X_{i+1} = X_i \cup \{A \in R \mid \text{van olyan } U \to V \in F, \text{ hogy } U \subseteq X_i \text{ és } A \in V\}$,
$\vdots$
$X^+(F) = X_{\text{utolsó}}$, (amikor már nem nõ)

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

Példa

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