Adatbázisok elmélete 16. előadás

Csima Judit
Budapesti Műszaki és Gazdaságtudományi Egyetem
Számítástudományi Tsz.
I. B. 136/b
csima@cs.bme.hu

2003. Április 8.

Hűséges felbontás két részre

Hogyan biztosíthatja F, hogy a felbontás hűséges legyen?

Tétel. *Az (R,F) séma $\rho = (R_1, R_2)$ felbontása hűséges $\iff$ vagy*
(a) $F \models R_1 \cap R_2 \to R_1 \setminus R_2$, *vagy*
(b) $F \models R_1 \cap R_2 \to R_2 \setminus R_1$.

Példa: R(TNÉV, TERMELŐ, ÁR, CÍM)
$F = \{$TERMELŐ $\to$ CÍM ; TNÉV, TERMELŐ $\to$ ÁR$\}$

$\rho = ($TERMELŐ, CÍM ;TNÉV, TERMELŐ, ÁR$)$
$R_1 \cap R_2 = \{$TERMELŐ$\}$, $R_1 \setminus R_2 = \{$CÍM$\}$, $R_2 \setminus R_1 = \{$TNÉV, ÁR$\}$
$\implies$ (TERMELŐ)$^+(F) = \{$TERMELŐ, CÍM$\} \supseteq R_1 \setminus R_2 \implies$ hűséges $\surd$

$\rho = ($TNÉV, TERMELŐ; TNÉV, CÍM, ÁR$)$
$R_1 \cap R_2 = \{$TNÉV$\}$, $R_1 \setminus R_2 = \{$TERMELŐ$\}$, $R_2 \setminus R_1 = \{$CÍM, ÁR$\}$
$\implies$ (TNÉV)$^+(F)$ nem tartalmazza egyiket sem $\implies$ nem hűséges ↯

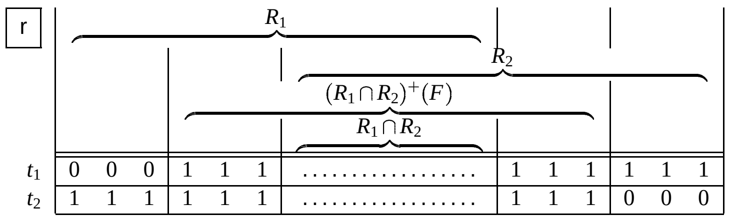
Bizonyítás: Tegyük fel, hogy $F \models R_1 \cap R_2 \to R_1 \setminus R_2$, belátjuk, hogy a felbontás hűséges. (Ha a másik igaz, ugyanígy.)

Legyen r egy tetszőleges reláció, $s = m_\rho(r)$. Elég belátni, hogy $s \subseteq r$, hiszen $r \subseteq s$ mindig igaz. Azaz, lássuk be, hogy ha t sora s-nek, akkor r-nek is.

Ha t sora s-nek, akkor $\exists u_1, u_2$ sorai r-nek, hogy $t[R_1] = u_1[R_1]$ és $t[R_2] = u_2[R_2]$.
$\implies u_1[R_1 \cap R_2] = t[R_1 \cap R_2] = u_2[R_1 \cap R_2]$
de ha két sor megegyezik a metszeten, akkor a feltétel miatt $R_1 \setminus R_2$-n is $\implies$ egyeznek az egész R_1-en $\implies u_2$ és t egyeznek R_1-en.
$\implies t = u_2$, hiszen R_1-en a fenti miatt, R_2-n a feltevés miatt egyeznek. $\surd$

Másik irány: Belátjuk, hogy ha $R_1 \setminus R_2 \not\subseteq (R_1 \cap R_2)^+(F)$ és $R_2 \setminus R_1 \not\subseteq (R_1 \cap R_2)^+(F)$, akkor ρ nem hűséges.

Legyen r a következő reláció:



A feltétel miatt a két szélső rész nem üres, ott nem egyezik meg a két sor.
r-ben igazak F függőségei (mint a teljességi tételnél).
Viszont $m_\rho(r) \supsetneq r$, hiszen $m_\rho(r)$-ben a csupa 1 sor is benne van. ↯

Hűségesség ellenőrzése általában

Adott (R,F) és $\rho = (R_1,\ldots,R_k)$, ahol $R = A_1,\ldots,A_n$.
Hogyan tudjuk eldönteni, hogy hűséges-e a felbontás?

Készítünk egy $k \times n$-es táblázatot:

	A_1	$\ldots$	A_j	$\ldots$	$A_{j'}$	$\ldots$	A_n
R_1							
$\vdots$							
R_i			a_j		$b_{ij'}$		
$\vdots$							
R_k							

- Kezdetben az (i,j) helyre a_j-t írunk, ha $A_j \in R_i$ és b_{ij}-t, ha $A_j \notin R_i$.

- Veszünk egy tetszőleges $X \to Y \in F$ függést.
  Ha két sor megegyezik X-en, akkor egyenlővé tesszük Y-on is az alábbi módon:
  - Ha valahol a_j és b_{ij} van, akkor a b_{ij}-t a_j-ra cseréljük.
  - Ha b_{kj} és b_{lj} van, akkor az egyiket átírjuk a másikra.
- Ezt minden függésre megcsináljuk tetszőleges sorrendben, szükség esetén többször is.

Tétel. ρ *pontosan akkor hűséges ha a végén lesz csupa a sor.*

Nem bizonyítjuk.

Példa a táblázatos tesztre

$R(ABCD) \quad F = \{A \to C;\ C \to B;\ B \to D\} \quad \rho = (AB, BC, ACD)$

	A	B	C	D
R_1	a_1	a_2	$b_{13} \to \boxed{a_3}$ [i]	b_{14}
R_2	b_{21}	a_2	a_3	b_{24}
R_3	a_1	$b_{32} \to \boxed{a_2}$ [ii]	a_3	a_4

[i] $A \to C$ miatt
[ii] $C \to B$ miatt

Lett csupa a sor $\Longrightarrow$ hűséges felbontás

Hűséges felbontás

Tétel. *Adott (R,F), $\rho = (R_1,\ldots,R_k)$ az R hűséges felbontása és $\sigma = (S_1,\ldots,S_m)$ az R_1 hűséges felbontása (azaz R_1-et tovább bontjuk). Ekkor $\tau = (S_1,\ldots,S_m,R_2,\ldots,R_k)$ hűséges felbontása R-nek.*

Bizonyítás: Legyen r egy R-re illeszkedő reláció és ennek R_1-re eső vetülete legyen r_1. Tovább bontva r_1-et σ szerint kapjuk az $s_1, s_2, \ldots, s_m$ vetületeket. Mivel σ hűséges, ezért $s_1 \bowtie s_2 \bowtie \cdots s_m = m_\sigma(r_1) = r_1$.
Mivel ρ is hűséges, ezért $r = m_\rho(r) = r_1 \bowtie r_2 \bowtie \cdots r_k$ Ebbe beírva r_1 helyére a σ hűségességéből kapott egyenlőséget, kapjuk, hogy $r = s_1 \bowtie s_2 \bowtie \cdots s_m \bowtie r_2 \bowtie \cdots r_k = m_\tau(r)$, azaz τ is hűséges. Itt persze használtuk $\bowtie$ asszociativitását.

Tétel. *Ha ρ hűséges és $\sigma \supseteq \rho$ (σ-ban több komponens van), akkor σ is hűséges.*

Bizonyítás: $r \subseteq m_\sigma(r) \subseteq m_\rho(r) = r$
A középső tartalmazás azért igaz, mert a keresztszorzatból szigorúbb feltételek szerint válogatunk.
$\Longrightarrow r = m_\sigma(r)$ $\surd$

Normálformák

Definíció. *Egy $X \to Y$ függés triviális, ha $Y \subseteq X$. (Mert ezek a függések nem hordoznak sok infót, mindg igazak.)*

Definíció (Boyce–Codd normálforma). *Az (R,F) relációs séma BCNF-ben van, ha tetszőleges nemtriviális $X \to A \in F^+$ függés esetén X szuperkulcs.*

Azaz csak olyan függések vannak, hogy a szuperkulcs mindent meghatároz.

Tétel. *Az (R,F) BCNF-ben van pontosan akkor ha tetszőleges $A \in R$-re és $X \subseteq R$ kulcsra igaz, hogy nincs olyan $Y \subseteq R$, amire $X \to Y \in F^+$; $Y \nrightarrow X$; $Y \to A \in F^+$ és $A \notin Y$. (Nincs tranzitív függés kulcstól.)*

Bizonyítás: Ha nincs BCNF-ben a séma, akkor van egy $Y \to A$ függés, ahol Y nem szuperkulcs és $A \notin Y$. Ekkor, tetszőleges X kulccsal: $X \to Y, Y \not\to X, Y \to A$, de $A \notin Y$, ami épp egy kulcstól való tramzitív függés.

Másrészt, ha van tranzitív függés kulcstól, azaz X olyan kulcs, amivel $X \to Y, Y \not\to$

$X, Y \to A$, de $A \notin Y$, akkor $Y \to A$ egy olyan függés, ami sérti a BCNF tulajdonságot, mert Y nem lehet szuperkulcs, ha $Y \not\to X$.

Miért jó a BCNF séma?
Ha $C \to B$; $B \to A$ teljesülne, de $B \to C$ nem, akkor ugyanaz a B érték több C érték mellett is előfordulhatna, de minden példánynál ugyanazt az A értéket is tároljuk $\implies$ **redundancia**.

Állítás. *≤ 2 attribútumos reláció mindig BCNF.*

Bizonyítás: Ha $A \to B \implies A$ kulcs. Ha $B \to A \implies B$ kulcs.

Mivel F^+-ra követeljük meg a feltételt, nehéz ellenőrizni.
DE:

Tétel. *Ha (R,F) nem BCNF, akkor van olyan $X \to Y \in \mathbf{F}$, amely jobboldalának valamely A attribútumára $X \to A$ nemtriviális és X nem szuperkulcs. (Az ilyen $X \to A \in F^+$.)*

Bizonyítás: Ha (R,F) nem BCNF, akkor van $U \to B \in F^+$, hogy U nem szuperkulcs és $B \notin U$. $\implies B \in U^+(F) \implies U \subsetneq U^+(F)$

Az algoritmus, ami $U^+(F)$-et számolja, el tud indulni $\implies \exists V \to W \in F$, melyre $V \subseteq U$, $W \nsubseteq U \implies V \to W$ jó lesz $X \to Y$-nak.
Ugyanis V nem szuperkulcs, hiszen $V \subseteq U$ és U nem szuperkulcs.
$W \nsubseteq U \implies \exists A \in W \setminus U \subseteq W \setminus V$, így $V \to W$ nem triviális.

Ez jelentősen könnyíti az ellenőrzést, csak F függőségeit kell végignézni, nem F^+-ét.