Adatbázisok elmélete 18. előadás

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

2003. Április 15.

3NF-re felbontás főtétele

Tétel. *Tetszőleges (R,F) sémának van hűséges és függőségőrző felbontása 3NF sémákra.*

Bizonyítását már láttuk a múlt órán, lényege az volt, hogy keresünk egy $G = \{X_1 \to A_1, \ldots, X_k \to A_k\}$ minimális fedést F-hez és egy X kulcsot és ekkor a jó felbontás $(X_1A_1, \cdots, X_kA_k, X)$ lesz.

Megjegyzés: Előfordulhat, hogy valamelyik X_iA_i már tartalmaz kulcsot. Ilyenkor a $\rho = \{X_1A_1, \ldots, X_kA_k\}$ is jó felbontás már, nem kell bevenni a kulcsot is.

Megjegyzés: Volt 3NF, lesz 4NF, mi van a többivel?
2NF már nem érdekes, 1NF kicsit érdekes, de nem foglalkozunk vele.

A tétel bizonyítása során felhasználtuk, hogy létezik minimális fedés. Ennek igazolása:
Állítás. *Tetszőleges F-nek van minimális fedése.*

Bizonyítás: Algoritmust adunk, külön gondoskodunk minden pont teljesítéséről.

(1) $X \to Y \in G,\ Y = A_1 \ldots A_k \implies$ minden $X \to A_i$-t beveszünk, ha $A_i \notin X$.

(2) Minden $X \to A \in G$ függésre kiszámoljuk $Y := X^+(G \setminus \{X \to A\})$-t. Ha $A \in Y$, akkor $X \to A$ elhagyható, különben nem.

(3) Ellenőrizni kell, hogy $X \to A$ baloldala minimális-e. X minden B elemére kiszámoljuk $Y := (X \setminus \{B\})^+(G)$-t. Ha $A \in Y$, akkor $X \to A$ helyett vegyük be $X \setminus \{B\} \to A$-t. Ha egyik B-re se lesz ilyen, akkor X minimális.

Megjegyzés: És persze a fenti három lépés során a függéshalmaz lezártja nem változik.

Példa minimális fedés kiszámolására

$R = (A,B,C,D) \qquad F = \{AB \to CD;\ AC \to BD;\ C \to A;\ C \to B\}$

(1) $F' = \{AB \to C;\ AB \to D;\ AC \to B;\ AC \to D;\ C \to A;\ C \to B\}$

(2) $C \to B$ miatt $AC \to B$ elhagyható és $AB \to C$ és $AC \to D$ miatt $AB \to D$ elhagyható, de más nem, ezt végig lehet nézni. $F'' = \{AB \to C;\ AC \to D;\ C \to A;\ C \to B\}$

(3) $C \to A$ miatt $AC \to D$ baloldaláról A elhagyható.
$F''' = \{AB \to C;\ C \to D;\ C \to A;\ C \to B\}$

Ez már minimális fedés.

A minimális fedés nem feltétlenül egyértelmű!

Példa: $R(A,B,C) \qquad F = \{AB \to C;\ A \to B;\ B \to A\}$ esetén jó minimális fedés lesz
$G_1 = \{B \to C;\ A \to B;\ B \to A\}$ és
$G_2 = \{A \to C;\ A \to B;\ B \to A\}$ is.

Példa: 3NF-re bontás

$R = (A,B,C,D,E) \qquad F = \{AE \to BC;\ AC \to D;\ CD \to BE;\ D \to E\}$

Ez nem 3NF, mert a kulcsok:
semelyik egyelemű halmaz nem kulcs (csak D lehetne, de az ő lezártja csak DE), viszont kételeműek közül szuperkulcs lesz AC, AD, AE (A-nak benne kell lennie minden kulcsban, mert A nincs jobboldalon), AB viszont nem szuperkulcs.
Ezek kulcsok is lesznek, mert egyik egyelemű se volt kulcs.
Más kulcs nincs is, mert ha lenne legalább háromelemű halmaz, aminek a lezártja az egész, akkor abban A biztos benne van és legalább C vagy D vagy E is benne van, de akkor az már csak szuperkulcs lehet, mert tartalmaz kulcsot.

Innen látszik, hogy a prímattribútumok: A, C, D, E, vagyis B nem az.

Tehát a $CD \to B$ függés rossz a 3NF szempontjából, mert CD nem szuperkulcs és B nem prím.

Csináljunk hát egy 3NF-ekre való függőségőrző, hűséges felbontást.

(1) $F' = \{AE \to B;\ AE \to C;\ AC \to D;\ CD \to B;\ CD \to E;\ D \to E\}$

(2) $AE \to C, AC \to D, CD \to B$ miatt $AE \to B$ elhagyható és $D \to E$ miatt $CD \to E$ is elhagyható, de más nem, ezt végig lehet nézni (mert például AE-nek a maradék függésekre vett lezártjában nincsen benne C).
$F'' = \{AE \to C;\ AC \to D;\ CD \to B;\ D \to E\}$

(3) Semelyik baloldal nem csökkenthető, mert például A lezártjában nincsen benne E, és a többi is ugyanígy látszik. Vagyis F'' már minimális fedés.

A minimális fedés alapján a jó felbontás:
(AEC, ACD, CDB, DE) mivel kulcsot nem is kellett hozzávennünk, mert az már benne van az egyik tagban (pl. AE az elsőben).