

Gagnasafnsfræði

Páll Melsted

30. sept

Uppbrot á töflu

Ef R er tafla með dálka \bar{A} þá getum við brotið R upp í töflur S og T með

1. $\bar{A} = \bar{B} \cup \bar{C}$
2. $S = \pi_B(R)$
3. $T = \pi_C(R)$

Sum uppbrot eru góð, t.d.

- {title,year,length,studioName},
- {title,year,starName}

En önnur eru slæm

- {title,year}
- {year,length,starName,studioName}

BCNF (Boyce-Codd Normal Form)

BCNF lýsir því hvernig má brjóta upp töflur til að losna við frávik (e. anomaly).

Vensl R eru á **BCNF** formi þþaa að ef

$$\bar{A} \rightarrow \bar{B}$$

gildir (og eru ófáfengileg, þ.e. $\bar{B} \subsetneq \bar{A}$)

þá er \bar{A} **yfirlykill**.

BCNF Reiknirit

Ef R er á BCNF formi þá skilum við $\{R\}$

Finnum einhver BCNF frávik, $X \rightarrow Y$

- Reiknum X^+
- Setjum $R_1 = \pi_{X^+}(R)$
- Setjum $R_2 = \pi_L(R)$
- þar sem L er sammengið af X og þeim dálkum í R sem eru ekki í.

Reiknum fallákveður fyrir varpanirnar R_1 og R_2 köllum þær S_1 , S_2

Reiknum endurkvæmt BCNF fyrir og skilum sammenginu af niðurstöðunum

Uppbrot

Með því að koma venslum yfir á BCNF þá viljum við

1. Losna við öll frávik
2. Halda sömu upplýsingum
3. Varðveita tengsl, þ.e. fallákveður

BCNF uppfyllir 1 og 2.

Sömu upplýsingar

Þegar við brjótum upp vensl þá er hægt að fá upphaflegu töfluna með náttúrulegri tengingu (natural join).

Stundum er hægt að fá meira en venslin R sem við byrjuðum með

Chase reikniritið notar uppbrotið og fallákveður til að finna hvort gögnin varðveitist nákvæmlega.

3-NF

Vensl R eru á 3-NF þpaa ef

$$\bar{A} \rightarrow \bar{B}$$

gildir þá gildir annaðhvort

1. \bar{A} er yfirlykill

2. það B_i sem er ekki í \bar{A} er í einhverjum lykli
(ekki endilega sama lykli fyrir öll B_i)

3-NF Reiknirit

Inntak: vensl R , fallákveður F

1.Finnum lággrunn G fyrir F

2.Fyrir öll $\bar{A} \rightarrow B$ í G , notum $\bar{A} \cup \{B\}$ sem dálka fyrir ný vensl

3.Ef engin af dálkunum í skrefi #2 mynda yfirlykil, þá bætum við við venslum með lykil sem dálka.