Andmebaasid, MTAT.03.264
6. loeng
Normaalkujud

Anne Villems

Eelmise nädala 1. ülesanne

Relatsioon kajastab muusikakooli õpilaste vabariigi aastapäeva kontserti, kus esinevad vaid üksiksolistid.

Kontserdikava (Teose nimi, Teose autor,

N A

esitaja eesnimi, esitaja perenimi, instrument)

E P I

<Partita G-moll GWV1004, J.S.Bach, Peeter, Kask, tšembalo>

Kui iga esineja esineb 1 kord: E,P $\rightarrow$I, A, N

Kui iga instrument on kavas 1 kord: I $\rightarrow$E,P,A,N

Ilma muude kitsendusteta: N,A $\rightarrow$ E,P,I

(iga pala kavas 1 kord)

2. ülesanne

Eksam (TEesnimi, Tperenimi, M.nr, kp, kood,nim, maht, hinne)

E P M K A N T H

M $\rightarrow$E,P

A $\rightarrow$N,T

M,A,K $\rightarrow$H

Võti (ainuke): M,A,K

Ei ole 2NF: R1(M,A,K,H); R2(M,E,P); R3(A,N,T)

Kus me oleme?

Meil on:

- Relatsiooni mõiste
- Relatsioonalgebra
- Funktsionaalsete sõltuvuse pere F ja tema sulund F^+
- Võti
- Oskame kontrollida, kas $X \rightarrow Y$ kasutades X^+ s.t. ilma F^+

Täna

- Osa materjali eelloengus
- Toome sisse sõltuvuste pere katte mõiste
- Dekompositsiooni mõiste
- Dekompositsiooni omadused
- Mitu normaalkuju: 1NF, 2NF ja 3NF(järelloengu videos)

Sõltuvuste pere

Meile ei meeldi algoritmid, mis vaatavad läbi F^+. Leidsime juba ühe algoritmi, mis vaatab läbi F. Kas saab F väiksemaks muuta?

Sõltuvuste pere F kate G

- Sõltuvuste ekvivalents:
  - Olgu sõltuvused pered F ja G
  - Me ütleme, et nad on ekvivalentsed, kui $F^+ = G^+$
  - e. ütleme, et G katab F (ja vastupidi).
- Kontrollimiseks on vaja **iga** $X \rightarrow Y \in F$ puhul kontrollida, kas ta kuulub G^+
  - Seda kontrollime, leides X^+ G järgi ja vaadates, kas $Y \in X^+$
  - Sama iga sõltuvuse korral $\in G$, kas $Y \in X^+$ F järgi.

Minimaalne kate

- Lemma: Mistahes F jaoks leidub selline kate, kus kõigi sõltuvuste paremal pool on 1 tunnus.
- Ütleme, et pere F on minimaalne kate, kui:
  1. Iga sõltuvuse parem pool sisaldab 1 tunnust.
  2. Ei leidu selliseid sõltuvusi, mille eemaldamisel saadav pere oleks ekvivalentne eelmisega.
  3. Ei leidu sõltuvusi, mille vasakust poolest saaks tunnuseid eemaldada nii, et pere oleks ekvivalentne eelmisega.
- Teoreem: iga pere F omab minimaalset katet.

Minimaalse katte näide

F:	Min.kate 1:	Min. kate 2:
$AB \rightarrow C$	$AB \rightarrow C$	$AB \rightarrow C$
$C \rightarrow A$	$C \rightarrow A$	$C \rightarrow A$
$BC \rightarrow D$	$BC \rightarrow D$	$BC \rightarrow D$
$ACD \rightarrow B$	$CD \rightarrow B$	$D \rightarrow E$
$D \rightarrow EG$	$D \rightarrow E$	$D \rightarrow G$
$BE \rightarrow C$	$D \rightarrow G$	$BE \rightarrow C$
$CG \rightarrow BD$	$BE \rightarrow C$	$CG \rightarrow B$
$CE \rightarrow AG$	$CG \rightarrow D$	$CE \rightarrow G$
	$CE \rightarrow G$	

Sõltuvalt eemaldamise jrk-st saame 2 erinevat katet.

Relatsioonide dekompositsioon

Relatsioonide dekompositsioon

- Tähistame relatsiooni R atribuutide hulka

$\{A_1, A_2, \ldots, A_n\}$ tähega U.

- Def. Relatsiooni R(U) dekompositsiooniks nimetatakse relatsiooni R asendamist relatsioonide hulgaga $\rho=\{R_1, R_2, \ldots, R_k\}$ selliselt, et $\mathrm{U} = \mathrm{U}_1 \bigcup \mathrm{U}_2 \bigcup \ldots \bigcup \mathrm{U}_\mathrm{k}$.
- Hulgad U_i võivad olla lõikuvad.

Milleks dekompositsioonid?

- See annab andmebaasi administraatorile võimaluse läbi viia salaplaan: R(U) asemel hoida andmebaasis relatsioonide hulka $R_1(U_1)$, $R_2(U_2)$,… , $R_k(U_k)$. Saame nende R_i-de seisundid R seisundist projektsiooniga.
- Salaplaan õnnestub, kui suudame R(U) seisundi taastada $R_1(U_1)$, $R_2(U_2)$,… , $R_k(U_k)$ seisunditest loomuliku ühendi operatsiooniga.
- Kas see alati on nii?
- Vaatame näidet!

Näide

- Olgu R(A, B, C) seisund järgmine:

A	B	C
1	2	2
1	2	1
1	3	1

- Olgu dekompositsioon ρ={R_1, R_2) selline, et R_1=(A, B) ja R_2=(B, C), siis on R_1 seisund ja R_2 seisund vastavalt:

A	B
1	2
1	3

A	C
1	2
1	1

Ja nende loomulik ühend (üle A)

A	B	C
1	2	2
1	2	1
1	3	2
1	3	1

Tulemuses on 4 rida!

Info läks kaotsi!

Kadudeta ühendi omadus

Def. Olgu antud R, F ja $\rho=\{R_1, R_2, \ldots, R_k\}$. Olgu r relatsiooni R mistahes seisund ja $r_1, r_2, \ldots r_k$ vastavalt $R_1, R_2, \ldots, R_k$ seisundid, siis me ütleme, et dekompositsioon on **kadudeta ühendi** omadusega F suhtes, kui

$r = \pi(r) \otimes \pi(r) \otimes \ldots \otimes \pi(r)$ $\otimes$ on siin loomuliku ühendi märk.

s.t. relatsiooni seisund on taastatav oma vastavatest projektsioonidest. □

- Eelmise näite dekompositsioon ei olnud seda!

Kadudeta ühendi kontrolli algoritm

Antud $R(A_1, A_2, \ldots, A_n)$, F ja $\rho=\{R_1, R_2, \ldots, R_k\}$.

Meetod: Ehitame tabeli n veeru ja k reaga, elemendiks ai, kui R_i-s esineb tunnus A_j, $b_{i,j}$ vastasel juhul.

Kordame: vaatame läbi tabeli read ja F. Kui leiame $X \to Y$ sellise, et kaks tabeli rida langevad kokku X osas, muudame Y nii, et nad langevad ka kokku (s.t. asendame $b_{i,j}$ a_i-ga või kui mõlemad on erinevad b-d, muudame indeksid kokkulangevaks)

Näide: **Hange(Hankija nimi, Aadress, Toode, Hind)**

- Kas Hankija(Firma nimi, Aadress) ja

Tarne (Firma nimi, Kaup, Hind) on kadudeta ühendi omadusega?

F: {Firma nimi $\rightarrow$ Aadress, Firma nimi, Toode $\rightarrow$ Hind }

Firma nimi	Aadress	Toode	Hind
a_1	a_2	$b_{1,3}$	$b_{1,4}$
a_1	$B_{2,2} \Rightarrow a_2$	a_3	a_4

See dekompositsioon on kadudeta ühendi omadusega.

Erijuht: k=2, s.t. R(U),F ja $\rho=(R_1(U1), R_2(U2))$

- Teoreem: k=2 korral annab $\rho=(R_1, R_2)$ kadudeta ühendi parajasti siis kui kehtib üks kahest funktsionaalsest sõltuvusest:

  $U_1 \cap U_2 \rightarrow U_1\text{-}U_2$ või

  $U_1 \cap U_2 \rightarrow U_2\text{-}U_1$ NB! Nõuab, et üks kuulub F^+

- Tõestus:

	$U_1 \cap U_2$	$U_1\text{-}U_2$	$U_2\text{-} U_1$
R1	a … a	a … a	b … b
R2	a … a	b … b	a … a

Normaalkujud

Teine normaalkuju

Second normal form

- Teatavasti võib relatsioonil olla palju võtmeid.
- Kui tunnus kuulub mingi võimaliku võtme tunnuste hulka, nimetame tunnust esmaseks (*primary*), vastasel juhul sekundaarseks (*secondary*).
- **Def: Teine normaalkuju:** Me ütleme, et relatsioon R on teises normaalkujus (2NF), kui mistahes võtme X ning mistahes sekundaarse tunnuse A jaoks ei leidu sellist võtme atribuutide pärisalamhulka $Y \subset X$, et kehtiks $Y \rightarrow A$.

Näide

Varustamine (Hankija nimi, Hankija aadress,
Hankija telefon, Toode, Toote hind)

Võti: {Hankija nimi, Toode}

F:

Hankija nimi, Toode $\rightarrow$ Hind

Hankija nimi $\rightarrow$ Hankija aadress, Hankija telefon

- Teine neist kahest rikub 2.NF tingimusi.
- See relatsioon ei ole teisel normaalkujul.

Relatsiooni viimine teisele normaalkujule

- Teisele normaalkujule viiakse relatsioon dekompositsiooni teel.
- Kuidas seda teha? Millised tunnused jätta alles lähterelatsiooni, millised viia teise relatsiooni?
- Kui meil relatsioonil R atribuutide hulgaga U rikub teise normaalkuju tingimust funktsionaalne sõltuvus $Y \rightarrow Z$, siis tuleb teha järgmine dekompositsioon:
- $\rho=\{R_1, R_2\}$ selliselt, et $R_1(U - Z)$ ja $R_2(YZ)$.
- Väide: Iga relatsiooni, mis rikub 2NF, saab viia 2NF kujule.

Näide

Varustamine (<u>Hankija nimi</u>, Hankija aadress,
Hankija telefon, **<u>Toode</u>,** Toote hind)

Võti: {Hankija nimi, Toode}

F: Hankija nimi, Toode $\rightarrow$ Hind

Hankija nimi $\rightarrow$ Hankija aadress, Hankija telefon

Y Z

Dekompositsioon:

Hankija (<u>Hankija nimi</u>, Hankija aadress, Hankija telefon)

Varustamine (<u>Hankija nimi</u>, **<u>Toode</u>,** Toote hind)

Eelloengu küsimused

- Võtme puudumine
- 2NF ühesus
- Kas ainult $\neg$2NF on halb?
- Kas mõnikord peab eirama NF nõudeid?
- Kas alati leidub 2NF?
- Kas on muid võimalusi 2NF saavutamiseks peale dekompositsiooni?

… küsimused 2.

- Ka mitte 2NF relatsioon saab olla „hea“?
- Kas 2NF kasvatab AB mahtu?
- Kas on olemas universaalne NF?
- Kas praktika kasutab mitte 2NF relatsioone?
- Kas saab lõhkuda 1 tunnusega relatsioonideks?

Järgmises videoloengus

- Kuhu kadus 1. normaalkuju
- 3. normaalkuju
- Kuidas viia relatsioon 3. normaalkujule