# Andmebaasid MTAT.03.264
# 9.loeng
# Relatsioonarvutus

Anne Villems

# Kontrollime eelmises loengus tehtud teisendusülesannet:

## Mitu relatsiooni teil oli?

1. 4 relatsiooni?
2. 5 relatsiooni?
3. 6 relatsiooni?
4. 7 relatsiooni?
5. Mind ei olnud eelmisel korral.

# Relatsioonid (1., 2., 3., 4., 5., 6. samm)

Osakond (Nimetus, Tänav, Maja, Linn, Indeks, Juhataja_IK)

Töötaja (IK, Eesnimi, Perenimi, Osakonna_nimetus)

Projekt (Proj_nimi, Maksumus, Tähtaeg, Osakonna_nimi)

Laps (Töötaja_IK, Nimi, Sünnikuup)

Osaleb (Töötaja_IK, Projekti_nimetus, Tunde)

Vahearuanne (Projekti_nimetus, Aruande_esitamise_kuup)

# Tänase loengu ülesanne punkti saamiseks:

9.Nädala plokist leiate ülesande (lahti kuni pühapäevani).

**Ülesandeks:** millistele mõistetele on veel vaja abimaterjale – selliseid, nagu on juba olemas relatsiooni ja funktsionaalsete seoste jaoks.

Üks võimalik lahend: mittemingisuguseid abimaterjale pole vaja, piisab loengumaterjali tekstidest ja slaididest – siis nii kirjutagegi.

# Sissejuhatus korteežarvutusse

- Seni tunneme 2 andmebaaside manipuleerimiskeelt (Database Manipulation Language e. DML): relatsioonalgebra ja SQL.
- Selleks, et aru saada keelte taga olevast teooriast, vajame veel ühte keelt: relatsioonarvutust.
- Relatsioonarvutusel on kaks versiooni: korteežarvutus ja doomenarvutus.

# Mis on korteež?

- Korteež on lõplik jada e. järjend.
- Relatsiooni seisundi element (e. tabeli rida) on korteež.
- Tavaliselt kasutatakse korteeži märkimiseks kolmnurksulge < ja >.
- Korteeži näide:

  <41310222718, Ene, Lepp, 51, 3275>
- Andmebaasides kasutame korteeže, mis vastavad relatsiooni kirjeldusele.

# Mis on doomen?

- Mõistet kasutatakse paljudes teooriates.
- Andmebaaside teoorias kasutatakse doomeni mõistet andmeelemendi võimalike väärtuste hulga tähenduses.
- Tabeli ühes veerus olevad andmed pärinevad kõik ühest doomenist – antud tunnuse võimalike väärtuste hulgast.
- Doomen väärtuste hulgana võib olla lõplik või lõpmatu (eksami hindel lõplik hulk võimalikke väärtusi: {A, B, C, D, E, F}, tunnusel „lapse kaal“ on võimalikke väärtusi lõpmata palju.

# Korteežmuutuja

- On muutuja, mille väärtuseks on korteež.
- Andmebaasides on korteežmuutuja seotud relatsiooniga, s.t. korteežmuutujal on tüüp (järk, elementidel igal oma tüüp), vastavalt relatsiooni kirjeldusele.
- Kirjutis R(t), kus R on relatsioon ja t on korteežmuutuja, väärtus on „tõene“, kui muutuja t väärtus on relatsiooni R antud seisundis olemasolev korteež (e. rida R seisundi tabelis).

# Kuidas ehitatakse üles formaalne keel?

Tuletame meelde: matemaatilise loogika valemi definitsioon (diskreetne matemaatika):

1. samm: aatomite defineerimine

DEFINITSIOON

- 1. Iga lausemuutuja on lausearvutuse valem.
- 2. Kui $\mathcal{A}$ on valem, siis $\neg \mathcal{A}$ on valem.
- 3. Kui $\mathcal{A}$ ja $\mathcal{B}$ on valemid, siis

$(\mathcal{A} \& \mathcal{B})$, $(\mathcal{A} \vee \mathcal{B})$, $(\mathcal{A} \supset \mathcal{B})$ ja $(\mathcal{A} \sim \mathcal{B})$ on valemid.

2.-3. samm – aatomitest keerulisema lause ehitamine

- 4. Valemiteks loetakse ainult neid avaldisi, mis on seda tingimuste 1. – 3. põhjal.

4. Muude lausete välistamine

# Korteežarvutuse defineerimine

- **Def. Avaldiseks** korteežarvutuses nimetatakse kirjutist kujul:

  $$\{ t \mid \psi (t)\}$$

  kus t on korteežmuutuja ja $\psi$ on valem, mis on konstrueeritud järgnevas esitatud reeglite kohaselt.

Tulemuseks on relatsioon korteežidest t.

Selgituseks: $\psi$ on kreeka tähestiku eelviimane täht, loetakse psii, kasutatakse siin üldkujul valemi tähistamiseks

# Aatomite defineerimine korteežarvutuses

1. **R(s)** kus R on relatsiooni nimi ja s on korteežmuutuja. Semantika (e. tähendus): kirjutis R(s) väidab, et korteež s esineb relatsioonis R antud seisundis r, s.t. $s \in r$ ;
2. **s[ i ] Θ u[ j ]** kus s ja u on korteežmuutujad ja **Θ** $\in \{=, \neq, <, \leq, \geq, >\}$; Semantika: korteeži s i-s element on seoses **Θ** korteeži u j-inda elemendiga. Elemente nummerdatakse alates 1-st.
3. **s[ i ] Θ a** ja **a Θ s [ i ]** kus a on konstant.

# Näited korteežarvutuse aatomitest

Olgu relatsioon Paber (Formaat, Pikkus, laius)

Näit. <A4, 297, 210>

Aatomid:

Paber (s)

s[2]>s[3]

s[2] = 148

# Kas on vaja tuua aatomite näiteid?

1. Ei ole vaja
2. Ei tea, kas on vaja
3. Mõned veel vist võiks
4. Jah, on vaja
5. Ma ei saa veel mitte millestki aru

# Loengu ülesanne

- Kirjutage sinna, mille kohta on vaja abimaterjale (korteež, korteežarvutuse aatomid, jne.)

# Olgu 2 relatsiooni: R1(A,B, C) ja R2(E, G)
# Millised neist on aatomid?

1. R1(t) & t[3]= „TÜ“
2. R2(u)
3. u[1]=u[2]
4. t[2]<u[1]
5. 7=7
6. 3>7
7. R1=R2

# Valem ja tema vabad muutujad

1. **Iga aatom on valem**. Kõik aatomites esinevad muutujad on vabad.
2. Kui $\psi_1$ ja $\psi_2$ on valemid, siis on valem ka $\boldsymbol{\psi_1 \& \psi_2}$ , $\boldsymbol{\psi_1 \vee \psi_2}$ ja $\neg \boldsymbol{\psi_1}$
   Muutujad on vabad või seotud vastavalt sellele, kas nad on vabad või seotud valemites $\psi_1$ ja $\psi_2$ .
3. Kui $\psi$ on valem, kus s on vaba muutuja, siis **($\exists$s) ($\psi$)** on valem. Selles valemis s on seotud, teised muutujad seotud/vabad nagu nad olid valemis $\psi$. Semantika näide: ($\exists$ s) R(s) väidab, et R seisund ei ole tühi.
4. Kui $\psi$ on valem, kus s on vaba muutuja, siis **($\forall$s) ($\psi$)** on valem. Saadud valemis s on seotud, teised - sõltuvalt oma staatusest valemis $\psi$. Semantika näide: ($\forall$ s) (s[2]=5) väidab, et R antud seisundis iga korteeži teise liikme väärtus on 5, e. vastava tabeli teine veerg koosneb ainult arvudest „5“.
5. Ei ole muid valemeid peale siindefineeritute.

# Keerulisemad valemid ja tehete prioriteedid

- Selguse huvides võib valemites kasutada sulge, operatsioonide prioriteedid on kahanevas järjestuses järgmised:

  $\Theta$, $\forall$ ja $\exists$ , $\neg$ , & , V

- **Avaldis relatsioonarvutuses on kujul**

  **{ t | $\psi$ (t)}, kus t on ainuke vaba muutuja.**

# Näited

1° $\{ t \mid R(t) \vee S(t)\}$

2° $\{ t \mid R(t) \,\&\, \neg S(t)\}$

3° $\{ t \mid (R(t) \,\&\, ( t[2]=5 \vee t[2]=4)\}$

4° $\{ t \mid R(t) \,\&\, (t[3]>210 \,\&\, t[4]>148)\}$

# Relatsioonalgebra ja korteežarvutuse vahekord

- Kas uus keel (korteežarvutus) on võimsam või vähemvõimas kui relatsioonalgebra?
- **Def.** Nimetame korteežarvutuse avaldist ekvivalentseks relatsioonalgebra valemiga, kui nendes on kasutatud samu relatsioone ja nende relatsioonide mistahes seisundite juures on vastavate avaldiste väärtuserelatsiooni seisund sama.
- **Väide:** Igale relatsioonalgebra operatsioonile saab vastavusse seada temaga ekvivalentse relatsioonarvutuse valemi.

# Millise tehnikaga peaks seda väidet tõestama?

1. Vastuväiteliselt?
2. Induktsiooniga?
3. Konstruktiivselt?

# Väide: igale rel. algebra operatsioonile saab ehitada ekvivalentse korteežarvutuse valemi

1. **R⋃S** $\{t|\ R(t) \vee S(t)\}$ kooskõla!
2. **R-S** $\{t|\ R(t) \,\&\, \neg S(t)\}$ kooskõla!
3. **R x S** $\{ t^{(n+m)} |\ \exists u^{(n)} \exists v^{(m)} (R(u) \,\&\, S(v) \,\&\, t[1]=u[1] \,\&\, \ldots \,\&\, t[n]=u[n] \,\&\, t[n+1]=v[1] \,\&\, \ldots \,\&\, t[n+m]=v[m])\}$

(jätkub)

# Konstruktsiooni järg:

**4.** $\boldsymbol{\pi_{i1, \ldots, ik}}$ **(R)** $\{t^{(k)} \mid \exists u\ (R(u)\ \&\ t[1]=u[i_1]\ \&\ \ldots\ \&\ t[k]=u[i_k]\ )\}$

**5.** $\boldsymbol{\sigma_F}$**(R)** $\{t \mid R(t)\ \&\ F'\ \}$

kus F’ saadakse avaldisest F asendades kõik atribuutide nimed vastavate t elementidega t[i]

# Miks otsekorrutisele ja projektsioonile vastavas lauses on kasutatud kvantoreid, aga teistes mitte?

1. Ka neis kahes pole kvantoreid tegelikult vajagi
2. Tulemuse tüüp erineb argumentide tüübist
3. Ka teistesse peaks kvantorid lisama
4. Neid kahte kasutatakse kõige rohkem

# Probleem

- Korteežarvutus on vähemalt sama võimas kui relatsioonalgebra.
- Kas temaga saab teha midagi, mida relatsioonalgebraga ei saa?
- Kuidas arvutada

$$\{ t \mid \neg R(t) \}$$

???

# Probleemi olemus

- Korteežarvutuses saab kirja panna valemeid, millele vastab (potentsiaalselt) lõpmatu relatsioon.
- Selliseid relatsioone me leida ei taha!
- Nad ei huvita meid ka!
- Kitsendame korteežarvutust, nimetame loodavat alam-korteežarvutust **ohutuks** korteežarvutuseks.

# Ohutu valemi konstrueerimine

- Konstrueerime hulga DOM ($\psi$)
- Andmebaasi relatsioonides on mitut tüüpi andmeelemente. Vaatame neid literaalidena, s.t. sümbolite jadana, mis seda väärtust esindab.
- Koostame literaalide hulga Dom ($\psi$) järgnevalt:
  - Paneme sinna kõigi kasutatud relatsiooni seisundite kõik väärtused literaalidena
  - Lisame kõigi valmis $\psi$ kasutatud konstantide literaalid.
- Hulk DOM ($\psi$) on lõplik.

# Ohutu korteežarvutus

Def. Korteežarvutuse avaldis $\{t \mid \psi(t)\}$ on **ohutu**, kui:

1. alati, kui t rahuldab valemit $\psi(t)$, kuulub iga t komponent hulka DOM($\psi$);
2. iga $\psi(t)$ alamavaldise $(\exists\ u)\ (\omega(u))$ korral: kui u rahuldab valemit $\omega(u)$, siis iga u komponent kuulub hulka DOM($\psi$),
3. iga $\psi(t)$ alamavaldis $(\forall\ u)(w(u))$ korral: kui mistahes u komponent <u>ei kuulu</u> hulka DOM(y), siis u rahuldab $\omega(u)$. (NB! $\forall u(\omega(u)) \equiv \neg\exists u(\neg\ \omega(u))$ )

# Ohutute korteežarvutuse valemite näited

Kui osavalem $\psi(t)$ on ohutu, siis on ohutud ka:

1. $\{t \mid R(t) \,\&\, \psi(t)\}$,
2. $\{t \mid R(t) \,\&\, \neg\, S(t)\}$,
3. $\{t \mid R(t) \,\&\, S(t)\}$
4. $\{t \mid R_1(t) \vee R_2(t) \vee \ldots \vee R_k(t) \,\&\, \psi(t)\}$

# Relatsioonalgebra ja ohutu korteežarvutuse seos ja erinevus

**Teoreem:** Kui E on relatsioonalgebra avaldis, siis leidub temaga ekvivalentne ohutu korteežarvutuse avaldis.

- Relatsioonalgebra on **operatsiooniline keel** – kirjeldame missuguste operatsioonidega vastus leida.
- Korteežarvutus on **deklaratiivne keel** – kirjeldame, mida tahame saada.

# Relatsioonarvutuse ja korteežarvutuse ajalooline roll

- Relatsioonalgebra lõi aluse nn. väga kõrge taseme keelele, kus tehti operatsioone tervete relatsioonidega ja vaja oli ainult 5 operatsiooni
- Korteežarvutus lõi aluse deklaratiivsetele keeltele, mille abil ehitati reaalsed päringukeeled
- Järgmises videoloengus: doomenarvutus ja tema seos korteežarvutuse ja relatsioonalgebraga.