

Tietokantojen perusteet, K 2004

Relaatioalgebra

- Relatiomalliin liittyy malli tietokannan käsittelystä
- Tietokannasta pitää pystyä hakemaan tietoa ja toisaalta tietokantaa on ylläpidettävä
- Tietokannan käsittelyn malli **relaatioalgebra** määrittelee operaatiot, joilla olemassaolevista relaatioista voidaan laskea uusia relaatioita

Tietokantojen perusteet, K 2004

Relaatioalgebra

- Kyselyt:
 - lähtökohtana tietokannan tila joukkona relaatioita
 - kyselyn tuloksena yksi relaatio
 - kysely on lauseke, joka määrittelee miten lähtörelaatioista tuotetaan tulosrelaatio soveltamalla relaatioalgebran operaatioita
 - vrt aritmeettinen laskukaava.
 - Lähtökohtana joukko lukuja
 - Laskukaava kertoo miten luvuista lasketaan tulos soveltamalla laskentaoperaatioita

Tietokantojen perusteet, K 2004

Relaatioalgebra

- Ylläpito:
 - Relaatioalgebrassa ei tarkastella tietokannan ylläpitoa
 - Voitaisiin ajatella että ylläpito palautetaan kyselyiksi
 - Lähtökohtana joukko relaatioita
 - Kyselyllä muodostetaan uusi relaatio, joka korvaa jonkin lähtöjoukkoon kuuluneen relaation
 - > tietokannan tila on siis muuttunut

Tietokantojen perusteet, K 2004

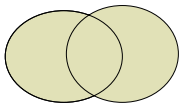
Relaatioalgebra

- Operaatiot, joilla relaatioista voidaan muodostaa uusia relaatioita
- Perustana joukko-oppi
- Joukko opin perusoperaatiot
 - yhdiste, erotus, ristitulo, leikkaus
- Erityisiä relaatioalgebran omia operaatioita
 - projekio, valinta, liitokset

Tietokantojen perusteet, K 2004

Relaatioalgebra

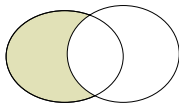
- Yhdiste (union)
 - Yhdisteen avulla muodostetaan relaatio joka sisältää kummankin yhdistettävän relaation monikot:
 - $R \cup S = \{t \mid t \in R \vee t \in S\}$,
 - missä R ja S ovat relaatioita ja t on relaation R tai S monikko.



Tietokantojen perusteet, K 2004

Relaatioalgebra - erotus (set difference)

- Erotuksessa tulosrelaatioon otetaan ne relaation monikot, jotka eivät sisälly erotettavaan relaatioon :
- $R - S = \{t \mid t \in R \wedge t \notin S\}$.



Tietokantojen perusteet, K 2004

Relaatioalgebra

- Yhdiste ja erotus edellyttävät, että relaatiot ovat samarakenteisia (union compatible)
 - sama määrä attribuutteja
 - vastinpaikoilla samat arvojoukot
 - vastinpaikkoihin liitettyjen attribuuttinimien ei tarvitse olla samoja
 - ensimmäinen osapuoli nimeää tuloksen sarakkeet

Tietokantojen perusteet, K 2004

Relaatioalgebra - ristitulo (Cartesian product)

- Ristitulossa $R \times S$ muodostetaan tulosrelaatioon monikoita, kokoamalla yhdeksi monikoksi arvot monikopareista, joissa parin monikoista ensimmäinen kuuluu relaatioon R ja toinen relaatioon S. Yhdistetty monikko muodostetaan jokaisesta monikoparista.

R	A	B
1	2	
3	4	

S	D	E
3	4	
5	6	
1	3	

$R \times S$	A	B	D	E
1	2	3	4	
1	2	5	6	
1	2	1	3	
3	4	3	4	
3	4	5	6	
3	4	1	3	

Tietokantojen perusteet, K 2004

Relaatioalgebra - ristitulo

- $aste(R \times S) = aste(R) + aste(S)$
- $koko(R \times S) = koko(R) * koko(S)$
 - opiskelija relaatiossa 30 000 monikkoa
 - opintosuoritus relaatiossa 600 000 monikkoa
 - opiskelija x opintosuoritus: 18 000 000 000
- Samannimiset attribuutit:
 - käytetään alkuperätarkennetta
 - $R(A,B,C) \times S(B,C,D) \Rightarrow R \times S(A, R.B, R.C, S.B, S.C, D)$

Tietokantojen perusteet, K 2004

Relaatioalgebra - projektiio (projection)

- Projektiiossa poimitaan relaatiosta annetuissa sarakkeissa esiintyvät arvot
- $\pi_{A_1, \dots, A_n}(R) = \{(a_1, \dots, a_n) \mid x \in R, \forall i=1..n: a_i = x.A_i\}$
 - A_1, \dots, A_n ovat attribuutteja
 - a_1, \dots, a_n ovat arvoja
 - x on monikko
 - $x.A$ on attribuutin A arvo monikossa x
- Vaikka sama arvoyhdistelmä a_1, \dots, a_n esiintyisi useassa lähtörelaation monikossa, se tulee kuitenkin tulosrelaatioon vain kertaalleen = toistuvien arvojen (dublikaattien) karsinta

Tietokantojen perusteet, K 2004

Relaatioalgebra - projektiio

T	A	B	D	E
1	2	3	4	
1	2	5	6	
1	2	1	3	
3	4	3	4	
3	4	5	6	
3	4	1	3	

$\pi_B(T)$	B
2	
4	

$\pi_{D,E}(T)$	D	E
3	4	
5	6	
1	3	

Tietokantojen perusteet, K 2004

Relaatioalgebra - valinta (selection)

- Valinnalla poimitaan ehdon täyttävät rivit
- $\sigma_{ehto}(R) = \{ x \mid x \in R \text{ ja } ehto \text{ on voimassa, kun siinä esiintyvät attribuutit korvataan niiden arvoilla monikossa } x \}$
- Ehdossa vertailtavina voivat olla attribuutit ja vakiot. Vertailuoperaattoreina tulevat kyseeseen $=, \neq, <, >, \leq$ ja \geq .

Tietokantojen perusteet, K 2004

Relaatioalgebra - valinta

R	A	B
1	2	
3	4	

$\sigma_{A=1}(R)$	A	B
1	2	

$\sigma_{3>2}(R)$	A	B
1	2	
3	4	

$\sigma_{A=5}(R)$	A	B

Tuloksena tyhjä joukko

Tietokantojen perusteet, K 2004

Relaatioalgebra -sijoitus

- Sijoituksella (assign) voidaan nimetä tulosrelaatio
- $S(A,B,\dots,N) :=$ lauseke.
- Lausekkeen tuloksen asteen täytyy olla sama kuin vasemmalla puolella olevien attribuuttien lukumäärä
- Sijoitusta ei yleensä pidetä relaatioalgebran operaationa, mutta sitä käyttäen voidaan nimetä välituloksia ja kyselyiden laadinta helpottuu
- Onimet(nimi):= $\pi_{\text{Sukunimi}}(\text{Opiskelija})$

Tietokantojen perusteet, K 2004

Relaatioalgebra - leikkaus

- Leikkaus on joukko-opin operaatio, jolla saadaan tulokseksi kahden joukon yhteiset alkiot. Se voidaan esittää erotus-operaation avulla.
- $R \cap S \equiv R - (R - S) \equiv S - (S - R)$

Tietokantojen perusteet, K 2004

Relaatioalgebra - loogiset lausekkeet ehdoissa

- Valintaehdoissa voidaan käyttää samanlaisia loogisia lausekkeita kuin ohjelmointikielissä, sillä niitä käyttävä valinta voidaan esittää myös yksinkertaisia valintoja ja yhdistettä, erotusta ja leikkausta käyttäen, esim.
- $\sigma_{\text{ehto1 or ehto2}}(R) \equiv \sigma_{\text{ehto1}}(R) \cup \sigma_{\text{ehto2}}(R)$ ja
- $\sigma_{\text{ehto1 and ehto2}}(R) \equiv \sigma_{\text{ehto1}}(R) \cap \sigma_{\text{ehto2}}(R)$
- $\sigma_{\text{not ehto1}}(R) \equiv R - \sigma_{\text{ehto1}}(R)$

Tietokantojen perusteet, K 2004

Relaatioalgebra -liitokset (join)

- Monikkoparin yhdistäminen jonkin ehdon perusteella
- Liitoksessa yhdistyy valinta ja ristitulo
 - $R \bowtie_{\text{liitosehto}} S \equiv \sigma_{\text{liitosehto}}(R \times S)$
- Yleisin liitos on yhdistää monikko ja siihen viittaava monikko
 - yhdistämisehtona on silloin ehto $R.A=S.VA$,
 - missä A on R:n avain ja VA on relaatioon R viittaava viiteavain S:ssä

Tietokantojen perusteet, K 2004

Liitos

R	A	B
1	2	
3	4	

S	D	E
3	4	
5	6	
1	3	

Ehto: $R.A=S.D$

RxS	A	B	D	E
1	2	3	4	
1	2	5	6	
1	2	1	3	
3	4	3	4	
3	4	5	6	
3	4	1	3	

Ensin ristitulo

Tietokantojen perusteet, K 2004

Liitos

R	A	B
1	2	
3	4	

S	D	E
3	4	
5	6	
1	3	

Ehto: $R.A=S.D$

RxS	A	B	D	E
1	2	3	4	
1	2	5	6	
1	2	1	3	
3	4	3	4	
3	4	5	6	
3	4	1	3	

Sitten karsinta

Tietokantojen perusteet, K 2004

Liitos

R	A	B
1	2	
3	4	

S	D	E
3	4	
5	6	
1	3	

Ehto: $R.A \neq S.D$

RxS	A	B	D	E
1	2	3	4	
1	2	5	6	
1	2	1	3	
3	4	3	4	
3	4	5	6	
3	4	1	3	

Tietokantojen perusteet, K 2004

Liitos

- Auto(Reknro, Väri, Vmalli, Merkki)
- Omistus(Reknro, Henkilönumero, Osoite, Nimi),

■ Autot, joilla ei ole omistajaa?

- Ei: $\pi_{Auto.reknro} (Auto \bowtie_{Auto.Reknro \neq Omistus.Reknro} Omistus)$

= $\pi_{Auto.reknro} (Auto \times Omistus - Auto \bowtie_{Auto.Reknro = Omistus.Reknro} Omistus)$

vaan

$\pi_{Auto.Reknro}(Auto) - \pi_{Reknro}(Omistus)$

Tietokantojen perusteet, K 2004

Liitos

Auto	Reknro	..
ABC-123		
DEF-456		
GHI-789		

Omistus	Reknro	Omistaja
	ABC-123	Liisa
	ABC-123	Kalle
	GHI-789	Pekka

Tulos	Reknro
	ABC-123
	DEF-456
	GHI-789

$\pi_{Auto.reknro} (Auto \bowtie_{Auto.Reknro \neq Omistus.Reknro} Omistus)$

Tietokantojen perusteet, K 2004

Luonnollinen liitos

- R*S
- liitosehto muodostetaan automaattisesti siten, että liitosehtona vaaditaan kaikkien vastinattribuuttien yhtäsuuruutta.
 - Vastinattribuutilla tarkoitetaan tässä sellaista attribuuttia, joka esiintyy kummassakin relaatioissa.
 - Edelleen, koska jokaisella vastinattribuutilla edellytetään olevan sama arvo kummassakin yhdistettävässä monikossa, attribuutti otetaan mukaan tulosrelaatioon vain kertaalleen.

Tietokantojen perusteet, K 2004

Luonnollinen liitos

- Olkoot A_1, \dots, A_n R:n attribuutit, jotka eivät esiinny S:ssä ja C_1, \dots, C_m S:n attribuutit, jotka eivät esiinny R:ssä sekä B_1, \dots, B_k attribuutteja, jotka esiintyvät kummassakin relaatiokaaviossa. Tällöin
- $R^*S \equiv \pi_{A_1, \dots, A_n, R.B_1, \dots, R.B_k, C_1, \dots, C_m} (R \bowtie_{R.B_1=S.B_1 \text{ and } \dots \text{ and } R.B_k=S.B_k} S)$

Tietokantojen perusteet, K 2004

Luonnollinen liitos

- $R(A,B,C) * S(A,D,E) = \pi_{A,B,C,D,E} (R \bowtie_{R.A=S.A} S)$
- $Q(A,B,C) * T(A,B,C) = \pi_{A,B,C} (Q \bowtie_{Q.A=T.A \text{ and } Q.B=T.B \text{ and } Q.C=T.C} T)$

Tietokantojen perusteet, K 2004

Ulkoliitos (outer join)

- Ulkoliitos on yhdisteen ja liitoksen yhdistelmä, jolla saadaan mukaan tulosrelaation myös sellaiset lähtörelaation monikot, joille liitosehdon mukaisesti ei löydy yhtään paria toisesta lähtörelaatiosta.
 - $R \bowtie_{\text{ehto}} S \equiv (R \bowtie_{\text{ehto}} S) \cup (R - \pi_{\text{att}(R)}(R \bowtie_{\text{ehto}} S)) \times \text{N}$
- $\text{att}(R)$ tarkoittaa kaikkia R:n attribuutteja ja N on yksimonikoinen relaatio, jonka kaavio on sama kuin relaatiolla S ja jonka jokainen arvo on tyhjäarvo.

Tietokantojen perusteet, K 2004

Ulkoliitos

- $\text{Auto} \bowtie_{\text{Auto.Reknro}=\text{Omistus.Reknro}} \text{Omistus}$
- Kaikki autot tulevat mukaan, mutta jos autolla ei ole omistajaa tulee kyseisen auton omistustiedoiksi tyhjäarvoja.

Tietokantojen perusteet, K 2004

Ulkoliitos

Auto	Reknro	..
ABC-123		
DEF-456		
GHI-789		

Omistus	Reknro	Omistaja
	ABC-123	Liisa
	ABC-123	Kalle
	GIH-789	Pekka

Tulos	A.Reknro	O.Reknro	Omistaja
	ABC-123	ABC-123	Liisa
	ABC-123	ABC-123	Kalle
	GIH-789	GIH-789	Pekka
	DEF-456	NULL	NULL