Prof. Dr. Ermir Rogova

Të dhënat e koorporatave dhe menaxhimi i databazave

Pjesa 6

Gjuhët Pyetëse (Query Languages)

- Si gjejmë informatat e ruajtura në një apo më shumë tabela? P.sh. si i gjejmë klientet që kanë llogari kursimi në nje degë të caktuar?

- Me anë të një gjuhe pyetëse !

- Janë dy kategori gjuhësh pyetëse:
  - Procedurale (Algjebra tabelore – Relational Algebra)
  - Declarative (Kalkulusi tableor – Relational Calculus, SQL)
- Gjuhët e "pastra":
  - Algjebra tabelore
  - Kalkulusi tabelor

- Gjuhët e pastra formojnë themelin e gjuhëve pyetësore të cilat janë në përdorim.

Algjebra tabelore

- Gjuhë procedurale

- Ka 6 operatorë themelor:
  - select: σ
  - project: $\prod$
  - union: $\cup$
  - set difference: –
  - Cartesian product: x
  - Rename ρ

- Operatorët marrin një apo më shumë tabela si të dhëna hyrëse dhe prodhojnë një tabelë të re si rezultat.

Operatori select

- Paraqitja: $\sigma_{p}(r)$
- *p* quhet **kallzuesi i përzgjedhjes**
- Definohet si:

$$\sigma_p(\boldsymbol{r}) = \{t \mid t \in r \text{ and } p(t)\}$$

Ku *p* është një formulë që parashtron kushtin përzgjedhës dhe përbëhet nga terma të lidhura me : $\wedge$ (**dhe**), $\vee$ (**ose**), $\neg$ (**jo**)

Çdo **term** është një nga:

<attribute> *op* <attribute> or <constant>

ku *op* është njëra nga: $=, \neq, >, \geq. <. \leq$

Operatori select– shembull

- Tabela r

A	B	C	D
$\alpha 1$	b	1	7
$\alpha 2$	b	5	7
$\alpha 3$	b	12	3
$\alpha 4$	b	23	10

- $\sigma_{D > 5 \wedge D < 8}(r)$

A	B	C	D
$\alpha 1$	b	1	7
$\alpha 2$	b	5	7

Operatori Project

- Paraqitja: $\prod_{A_1, A_2, \ldots, A_k}(r)$

  Ku A_1, A_2 janë emra atributesh (kolonash) dhe r është emer tabele.

- Rezultati definohet si tabelë me k kolona e cila përfitohet duke fshirë kolonat të cilat nuk janë në listë

- Rreshtat duplikatë poashtu largohen nga rezultati, pasiqë tabelat përfaqsojnë bashkësi.

Operatori Project– Shembull

Tabela *r*:

A	B	C
$\alpha 1$	b	1
$\alpha 2$	b	1
$\alpha 3$	b	1
$\alpha 4$	b	2

$\prod_{B,C} (r)$

B	C
b	1
b	1
b	1
b	2

=

B	C
b	1
b	2

Operatori Union

- Paraqitja: $r \cup s$

- Definohet si:

$$r \cup s = \{t \mid t \in r \text{ or } t \in s\}$$

- Që $r \cup s$ të jetë valid;
  1. *r, s* duhet të kenë të njëjtin **aritet** (të njëjtin numër të kolonave)
  2. Domenet e kolonave duhet të jenë **kompatibil** (p.sh. kolona e parë e *r* ka të njëjtin domen si kolona e parë e *s*, *etj*.)

Operatori Union – Shembull

Tabelat *r, s:*

A	B
α	1
α	2

r

A	B
α	2
α	3

s

$r \cup s$:

A	B
α	1
α	2
α	3

Operatori ndryshimi (Set Difference)

- Paraqitja: $r - s$

- Definohet si:

$$r - s = \{t \mid t \in r \textbf{ and } t \notin s\}$$

- Ndryshimi duhet të merret në mes të tabelave **kompatibile**.
  - *r* dhe *s* duhet të ketë të njëjtin aritet
  - Domenet e kolonave *r* dhe *s* duhet të jenë kompatibile

Operatori ndryshimi – shembull

Tabelat *r*, *s*:

A	B
α	1
α	2

r

A	B
α	2

s

$r - s$:

A	B
α	1

$s - r$?

Operatori produkti kartezian

- Paraqitja: $r \times s$

- Definohet si:

  $r \times s = \{t\, q \mid t \in r \textbf{ and } q \in s\}$

- Supozojmë që kolonat e r(R) dhe s(S) nuk kanë prerje. (dmth, $R \cap S = \emptyset$).

- Nëse kolonat e *r(R)* dhe *s(S)* kanë prerje, atëherë duhet të riemërohen.

Produkti kartezian – shembull

Tabelat *r*, *s*:

r

A	B
$\alpha 1$	1
$\alpha 2$	2

s

C	D	E
c1	10	e
c2	10	e
c3	20	e

r x *s*:

A	B	C	D	E
$\alpha 1$	1	c1	10	e
$\alpha 1$	1	c2	10	e
$\alpha 1$	1	c3	20	e
$\alpha 2$	2	c1	10	e
$\alpha 2$	2	c2	10	e
$\alpha 2$	2	c3	20	e

Operatori rename

- Lejon riemrimin dhe me kete edhe referimin e rezultateve të shprehjeve në algjebrën tabelore

- Mundeson referimin e një tabele me një emër tjetër.

- Shembull:

$$\rho_{x}(E)$$

 kthen shprehjen E me emrin X
- Nëse një shprehje algjebrike *E ka aritetin n*, atëherë

$$\rho_{x(A1,a2,\ldots,An)}(E)$$

- Kthen rezultatin e shprehjes E me emrin X dhe me atributet e riemruara A1, A2, …, An

Kombinimi i operatorëve

- Mund të formojmë shprehje duke përdorur disa operatorë.

- Shembull: $\sigma_{D='10'}(r \times s)$

$r \times s$

A	B	C	D	E
$\alpha 1$	1	c1	10	e
$\alpha 1$	1	c2	10	e
$\alpha 1$	1	c3	20	e
$\alpha 2$	2	c1	10	e
$\alpha 2$	2	c2	10	e
$\alpha 2$	2	c3	20	e

$\sigma_{D='10'}(r \times s)$

A	B	C	D	E
$\alpha 1$	1	c1	10	e
$\alpha 1$	1	c2	10	e
$\alpha 2$	2	c1	10	e
$\alpha 2$	2	c2	10	e

Operatorët shtesë

- Janë operatorë të cilët nuk fuqizojnë më tej algjebrën tabelore por të cilët i thjeshtojnë pyetjet e shpeshta.
- Prerja
- Bashkimi natyral (Natural join)

Operatori i prerjes (Set-Intersection)

- Paraqitja: $r \cap s$

- Definohet si:

  $$r \cap s = \{ t \mid t \in r \textbf{ and } t \in s \}$$

- Supozojmë që:
  - r, s kanë të njëjtin aritet
  - Kolonat e r dhe s janë kompatibil

- Mbani mend: $r \cap s = r - (r - s)$

Operatori i prerjes– shembull

Tabelat *r, s*:

A	B
α1	1
α2	2

r

A	B
α2	2

s

$r \cap s$

A	B
α2	2

Operatori bashkimi natural

- Paraqitja: $r \bowtie s$

- Nëse *r* and *s* janë tabela me skema respektive *R* and *S*, atëherë $r \bowtie s$ është një tabelë me skemën $R \cup S$, e cila përfitohet si më poshtë:
  - Konsiderojmë çiftet e rreshtave t_r prej *r* dhe t_s prej *s*.
  - nëse t_r dhe t_s kanë të njëjtën vlerë në secilën nga kolonat në $R \cap S$, shtpjmë në rezultat një rresht *t* ku
    - *t* ka të njëjtat vlera si t_r në *r*
    - *t* ka të njëjtat vlera si t_s në *s*

- Shembull:

  $r = (A, B, C, D)$

  $s = (B, D, E)$
  - Skema rezultuese = (A, B, C, D, E)
  - $r \bowtie s$ definohet si:

$$\prod_{r.A,\ r.B,\ r.C,\ r.D,\ s.E} (\sigma_{r.B = s.B \wedge r.D = s.D} (r \times s))$$

Operatori bashkimi natural– Shembull

Tabelat r, s:

A	B	C	D
$\alpha 1$	1	c1	d1
$\alpha 2$	2	c2	d2

r

B	D	E
1	d1	e1
2	d3	e2

s

$r \bowtie s$

A	B	C	D	E
$\alpha 1$	1	c1	d1	e1

Funksionet dhe operacionet grumbulluese (agregate)

- **Funksioni grumbullues** merr një koleksion vlerash dhe kthen si rezultat një vlerë të vetme:

  avg: vlera mesatare
  min: vlera minimale
  max: vlera maksimale
  sum: shuma e vlerave
  count: numri i vlerave

- **Operacioni grumbullues** në algjebrën tabelore:

$$_{G_1, G_2, \ldots, G_n}\vartheta_{F_1(A_1), F_2(A_2, \ldots, F_n(A_n)}(E)$$

E është çfardo shprehje në algjebër tabelore
  - $G_1, G_2 \ldots, G_n$ është listë e kolonave të cilat grupohen (mund të jetë e zbrazët)
  - Secili F_i është funksion grumbullues
  - Secila A_i është emër kolone

Operacioni grumbullues – Shembull

Tabela *r*:

A	B	C
α	α	7
α	β	7
β	β	3
β	β	10

$g_{\textbf{sum(c)}}(\mathrm{r})$

sum(c)
27

Operacioni grumbullues– shembull 2

Tabela *account* e grupuar sipas *branch-name*:

branch_name	*account_number*	*balance*
Perryridge	A-102	400
Perryridge	A-201	900
Brighton	A-217	750
Brighton	A-215	750
Redwood	A-222	700

$$_{branch\_name}\ \mathcal{G}\ _{\mathbf{sum}(balance)}\ (account)$$

branch_name	**sum**(*balance*)
Perryridge	1300
Brighton	1500
Redwood	700

Pyetje ???