



# Četvrta Nedelja – Optimizacija upita –

**Autor: Miloš Cvetanović**

# Interpretacija upita

- **Upotreba upitnih jezika (nezavisni od aplikativnog nivoa)**
- **Veliki broj ekvivalentnih upita**
- **Parsiranje upita**
- **Sintaksna provera upita (SQL sintaksa)**
- **Semantička provera (...postojanje odgovarajućih relacija)**
- **Interna forma upita (relaciona algebra)**
  - Identifikacija podupita
- **Traženje ekvivalentnih upita (prva faza)**
  - Obrada operacije selekcije
  - Obrada operacije projekcije
  - Obrada operacije prirodnog spajanja
  - Obrada ostalih operacija
- **Selekcija mogućih strategija izvršavanja (druga faza)**
  - Upotreba indeksa
  - Redosled obrađivanja zapisa
  - Procena broja pristupa disku

# Ekvivalentni upiti 1

- **AUTOR(SifA, Ime, Adresa)**  
**NASLOV(SifN, Naziv, SifO)**  
**JE\_AUTOR(SifA, SifN, Koji)**
- $\Pi_{SifA, Ime}(\sigma_{SifO="PJ"}(\text{naslov } x_* \text{ je\_autor } x_* \text{ autor}))$
- $\Pi_{SifA, Ime}(\sigma_{SifO="PJ"}(\text{naslov}) x_* \text{ je\_autor } x_* \text{ autor})$
- **PRAVILO: Izvršiti operaciju restrikcije što ranije!**
- $\Pi_{SifA, Ime}(\sigma_{SifO="PJ"} \& Koji > 1(\text{naslov } x_* \text{ je\_autor } x_* \text{ autor}))$
- $\Pi_{SifA, Ime}(\sigma_{SifO="PJ"} \& Koji > 1(\text{naslov } x_* \text{ je\_autor}) x_* \text{ autor})$
- $\Pi_{SifA, Ime}((\sigma_{SifO="PJ"}(\text{naslov})) x_* (\sigma_{Koji > 1}(\text{je\_autor})) x_* \text{ autor})$
- **PRAVILO:  $\sigma_{P1 \& P2}(e) = \sigma_{P1}(\sigma_{P2}(e)) = \sigma_{P2}(\sigma_{P1}(e))$**
- $(r1 x_* r2) x_* r3 = r1 x_* (r2 x_* r3)$   
 $r1 x_* r2 = r2 x_* r1$
- **PRAVILO: Redosled izvršavanja prirodnog spajanja je bitan!**

## Ekvivalentni upiti 2

- **PRAVILO: Izvršiti operaciju projekcije što ranije!**
- $\pi_{\text{SifA, Ime}}(\sigma_{\text{SifO}=\text{"PJ"}}(\text{naslov}) \times_{*} \text{je\_autor} \times_{*} \text{autor})$
- $\pi_{\text{SifA}}(\pi_{\text{SifN}}(\sigma_{\text{SifO}=\text{"PJ"}}(\text{naslov})) \times_{*} \text{je\_autor})) \times_{*} \pi_{\text{SifA, Ime}}(\text{autor})$
  
- **Potrebni atributi**
  - Oni koji se pojavljuju kao rezultat upita
  - Oni koji su potrebni za procesiranje upita
  
- **PRAVILA:**
  - $\sigma_p(r1 \cup r2) = \sigma_p(r1) \cup \sigma_p(r2)$
  - $\sigma_p(r1 - r2) = \sigma_p(r1) - r2 = \sigma_p(r1) - \sigma_p(r2)$
  - $(r1 \cup r2) \cup r3 = r1 \cup (r2 \cup r3)$
  - $r1 \cup r2 = r2 \cup r1$

# Procena cene izvršavanja upita

- $n_r$  broj zapisa u relaciji r
- $S_r$  veličina zapisa relacije r
- $V(A, r)$  broj različitih vrednosti atributa A relacije r
  
- $\sigma_{A=a}(r) \rightarrow n_r/V(A, r)$  uniformna raspodela vrednosti
  
- $r1(R1), r2(R2)$
- $r1 \times_* r2 = r1 \times r2$  R1, R2 nemaju zajedničkih atributa
- $n(r1 \times_* r2) < n_{r2}$  R1, R2 imaju zajedničkih atributa (ključ od R1)
- $n_{r1}n_{r2}/V(A, r_2)$  R1, R2 imaju zajednički atribut atribut A
- $n_{r1}n_{r2}/V(A, r_1)$  Uloga u spajanju je od značaja!
  
- **Indices & Hash**      **Clustering**
- **B+ stabla**
- **Strategije spajanja**
  - Jednostavna iteracija
  - Blok orjentisano spajanje
  - Strategija sjedinjavanja
  - Upotreba privremenih indeksa
  - Trostruko spajanje (istovremeno spajanje)