

Relačná algebra

1) RA

2) Operátory

3) Príklady

4) Vlastnosti operátorov

1) RA

Vieme, že **algebra** je matematický systém, ktorý sa skladá z **operandov** a **operátorov** (operácií). Operand je hodnota alebo premenná. Viac operandov spájame pomocou operátorov do **výrazu** (termu), novej hodnoty, nového operandu. Operátor je proces, ktorý z jedného alebo viac operandov vytvorí nový.

Jednotlivé algebry sa líšia s množinou hodnôt a typmi operátorov.

Definícia 1

k-členná **relácia** r nad množinami D_1, \dots, D_k je ľubovoľná podmnožina karteziánskeho súčinu,

$$r \subseteq D_1 \times \dots \times D_k,$$

kde D_i je doména i -tého atribútu a_i , teda množina všetkých prípustných hodnôt a_i .

Pripomíname, že relačná **schéma** R sme definovali ako

$$R(a_1:D_1, \dots, a_k:D_k),$$

kde

- R je názov relácie,
- a_j je názov atribútu,
- D_j je typ domény atribútu a_i .

Relačná **schéma** korešponduje s entitou a inštancia schémy s **reláciou** r , tabuľkou.

Definícia 2

Relačná algebra (RA) je také algebra

- operandy ktorej sú **relácie** alebo premenné, reprezentujúce relácie a
- operátory ktorej podporujú rôzne manipulácie s reláciami v databáze (napr. σ, π, \bowtie).

Relačná algebra je uzavretý systém – nielen operandy sú reláciami, ale aj výsledky operácií.

Relačná algebra je jedným z dopytovacích jazykov, ako je aj SQL alebo relačný kalkulus. Relačné DB systémy používajú RA na **optimalizáciu** dopytov.

2) Operátory

Operátory delíme na **základné** a **dopĺňujúce**, ktoré boli dodefinované neskoršie.

Základné operátory

1. Selekcia/Selection – horizontálna filtrácia σ
2. Projekcia/Projection – vertikálna filtrácia π
3. Premenovanie atribútov ρ
4. Karteziánsky súčin, Theta a Natural Join $\times, \bowtie_{\phi}, \bowtie$
5. Množinové operácie - Union, Intersection, Difference

Dopĺňujúce operácie

- Agregácie - SUM, AVG, COUNT, MIN, and MAX
- Outer Join

1. Selekcia – horizontálna filtrácia

Selekcia s podmienkou/výrokovou formou ϕ : $r_2 = \sigma_{\phi}(r_1)$

Návštevy u lekára s id k

$\sigma_{idL=k}(\text{Navstevy}) \Leftrightarrow \text{SELECT } * \text{ FROM Navstevy WHERE idL=k}$

2. Projekcia – vertikálna filtrácia

Projekcia s atribútm a1, a2: $r_2 = \pi_{a1,a2}(r_1)$

Mená a špecializácie lekárov

$\pi_{\text{meno,spec}}(\text{Lekari}) \Leftrightarrow \text{SELECT } \text{meno, spec} \text{ FROM lekari}$

3. Premenovanie atribútov

$\rho_{\text{Meno} / \text{NewMeno}}(r)$

$\rho_{\text{mesPrijem} / \text{prijem}}(\text{pacienti}) \Leftrightarrow \text{SELECT } \text{mesPrijem AS prijem} \text{ FROM pacienti}$

4. Karteziánsky súčin, Theta a Natural Join

Karteziánsky súčin: $r_3 = r_1 \times r_2$

$r_1 \times r_2 \Leftrightarrow \text{SELECT } * \text{ FROM } r_1, r_2$

Theta Join s podmienkou φ : $r_3 = r_1 \bowtie_{\varphi} r_2$

$$r_1 \bowtie_{\varphi} r_2 \Leftrightarrow \sigma_{\varphi}(r_1 \times r_2)$$

\Leftrightarrow SELECT * FROM r1, r2 WHERE φ

\Leftrightarrow SELECT * FROM r1 JOIN r2 ON φ

Natural Join: $r_3 = r_1 \bowtie r_2$

$$r_1 \bowtie r_2 \Leftrightarrow r_1 \bowtie_{r_1.id=r_2.id} r_2$$

\Leftrightarrow SELECT * FROM r1 JOIN r2 ON r1.id = r2.id

Výhodou natural join je skrátený zápis bez uvedenia zodpovedajúcich spojovacích atribútov.

3) Príklady

a) Návštevy u lekára s id k doplnené o informácie o pacientoch.

Odôvodnite, ktoré riešenie je efektívnejšie:

$$\sigma_{idL=k}(\text{Navstevy} \bowtie \text{Pacienti}) \Leftrightarrow \sigma_{idL=k}(\text{Navstevy}) \bowtie \text{Pacienti}$$

\Leftrightarrow SELECT * FROM navstevy n JOIN Pacienti p ON n.idP=p.idP
WHERE n.idL=3

\Leftrightarrow SELECT * FROM (SELECT * FROM Navstevy WHERE idL=3) n
JOIN Pacienti p ON n.idP=p.idP

b) Mená pacientov, ktorí navštívili lekára s id k:

$$\pi_{\text{meno}}(\sigma_{idL=k}(\text{Navstevy} \bowtie \text{Pacienti}))$$

$$\Leftrightarrow \pi_{\text{meno}}(\sigma_{idL=k}(\text{Navstevy}) \bowtie \text{Pacienti})$$

$$\Leftrightarrow \pi_{\text{meno}}(\sigma_{idL=k}(\pi_{idP}(\text{Navstevy})) \bowtie \text{Pacienti})$$

\Leftrightarrow
SELECT p.krstne FROM navstevy n JOIN Pacienti p ON n.idP=p.idP
WHERE n.idL=3

\Leftrightarrow
SELECT p.krstne FROM (SELECT * FROM Navstevy WHERE idL=3) n
JOIN Pacienti p ON n.idP=p.idP

\Leftrightarrow
SELECT p.krstne FROM (SELECT idP FROM Navstevy WHERE idL=3) n
JOIN Pacienti p ON n.idP=p.idP

c) Mená pacientov, ktorí navštívili kardiológa:

$$\pi_{\text{Pac.meno}} (\sigma_{\text{spec=kard}} (\text{Lekari} \bowtie \text{Navstevy} \bowtie \text{Pacienti}))$$

$$\Leftrightarrow \pi_{\text{meno}} \left(\pi_{\text{idL}} (\sigma_{\text{spec=kard}} (\text{Lekari})) \bowtie (\text{Navstevy} \bowtie \text{Pacienti}) \right)$$

$$\Leftrightarrow \pi_{\text{meno}} \left(\pi_{\text{idP}} \left(\pi_{\text{idL}} (\sigma_{\text{spec=kard}} (\text{Lekari})) \bowtie \text{Navstevy} \right) \bowtie \text{Pacienti} \right)$$

\Leftrightarrow

```
SELECT p.krstne FROM Lekari l
      JOIN Navstevy n ON n.idL=l.idL
      JOIN Pacienti p ON n.idP=p.idP
WHERE l.spec='kardiolog'
```

\Leftrightarrow

```
SELECT p.krstne FROM (SELECT idL FROM Lekari l WHERE spec='kardiolog') l
      JOIN Navstevy n ON n.idL=l.idL
      JOIN Pacienti p ON n.idP=p.idP
```

```
SELECT p.krstne FROM (SELECT idL FROM Lekari l WHERE spec='kardiolog') l
      JOIN (SELECT idL, idP FROM Navstevy) n ON n.idL=l.idL
      JOIN Pacienti p ON n.idP=p.idP
```

```
SELECT p.krstne FROM (SELECT idP FROM Lekari l, Navstevy n
      WHERE spec='kardiolog' and n.idL=l.idL) ln
      JOIN Pacienti p ON ln.idP=p.idP
```

3) Vlastnosti operátorov

Relačné operatory majú niektoré vlastnosti, ktoré sa využívajú optimalizátorom RDBS.

Selekcia σ je

- idempotentná

$$\sigma_{\phi}(\sigma_{\phi}(r)) = \sigma_{\phi}(r)$$

- komutatívna

$$\sigma_{\alpha}(\sigma_{\beta}(r)) = \sigma_{\beta}(\sigma_{\alpha}(r)).$$

Projekcia π

- je idempotentná

$$\pi_a(\pi_a(r)) = \pi_a(r)$$

- nie je komutatívna

$$\pi_a(\pi_b(r)) \neq \pi_b(\pi_a(r)).$$

Uveďte príklad na ilustráciu nekomutatívnosti projekcie.

Selekcia σ a π

- Selekcia a projekcia sú komutatívne

$$\sigma_{\alpha}(\pi_b(r)) = \pi_b(\sigma_{\alpha}(r))$$

práve vtedy, ak atribúty používané v podmienkach selekcie α tvoria podmnožinu atribútov b v projekcii. Vykonalenie selekcie pred projekciou môže byť užitočné, ak operand r je typu join.