Lista 2

Weronika Jakimowicz

07.03.2024

Zadanie 1.
Rozważmy relację R(A, B, C). Napisz zapytanie algebry relacji oraz zapytanie rrd/rrk, które zwróci pusty wynik wtedy i tylko wtedy gdy para atrybutów A, B jest kluczem relacji R.

Rozwiązanie.
ALGEBRA RELACJI

Pytamy, kiedy to co jest w kolumnach A i B nie jest unikalne dla danej krotki (nie jest kluczem). Ilość krotek mających daną parę w kolumnach A i B liczymy zapytaniem `gamma A, B, C; count C -> nr (R)`. Chcemy zwracać pusty wynik, kiedy w kolumnie nr widzimy 1, czyli wystarczy zrobić `sigma nr>1 (...)` i mamy gotowy wynik. W całości zapytanie algebry relacji prezentuje się następująco:

```
sigma nr > 1 (
  gamma A, B, C; count C -> nr (R)
)
```

RRK

W tym przypadku mamy troszkę szybsze rozwiązanie, bo od razu możemy spytać, czy istnieje inny element który spełnia pewne warunki:

$$\{x \in R \;:\; (\exists\ y \in R \setminus \{x\})\ y.A = x.A \;\wedge\; y.B = x.B \;\wedge\; y.C \neq x.C\}$$

RRD

Tutaj jest jeszcze szybciej, ale wolałam RRK

$$\{a, b, c \;:\; R(a, b, c) \;\wedge\; (\exists c')\ c \neq c' \;\wedge\; R(a, b, c')\}$$

Zadanie 2.
Rozważmy relację R(A, B, C) oraz S(X, Z), przy czym atrybut A jest kluczem w R. Napisz zapytanie algebry relacji oraz zapytanie rrk/rrd, które zwróci pusty wynik wtedy i tylko wtedy, gdy atrybut Z relacji S jest kluczem obcym wskazującym na atrybut A relacji R.

Rozwiązanie.
ALGEBRA RELACJI

Klucz obcy, to np. indeks studenta w tabeli ocen - wskazuje on wtedy na osobę w tabeli aktywnych studentów uniwersytetu, ale może się powtarzać w tabeli ocen.

Możemy zacząć od znalezienia krotek, które mają ten sam element w kolumnie A i kolumnie S przy pomocy joina: `R join A=Z S`. Nas interesują wszystkie te wpisy z relacji S, w których ta równość nie zachodzi. Rzutujemy więc wynik `join` na kolumny X i Z i odejmujemy wynik od S: `S - (pi X, Z (...))`. W całości dostajemy

```
S - (pi X, Z (
    R join A=Z S
    )
)
```

RRK

Wystarczy sprawdzić, czy nie istnieje element w R, który zgadza się z aktualnym elementem na kolumnie Z. Można skorzystać z praw de Morgana

$$\{x \in S \ : \ \neg(\exists y \in R)\ x.Z = y.A\} = \{x \in S \ : \ (\forall\ y \in R)\ x.Z \neq y.A\}$$

RRD

Tutaj nieco mniej elegancko jest zapisać przy pomocy $\forall$ moim zdaniem:

$$\{x, z \ : \ S(x, z) \ \wedge \ \neg[(\exists a, b, c)\ R(a, b, c) \ \wedge \ a = z]\} =$$
$$= \{x, z \ : \ S(x, z) \ \wedge \ (\forall\ a, b, c)\ \neg R(a, b, c) \vee a \neq z\}$$

Zadanie 3.

Dane są relacje R, S i T o schematach $R = AB$, $S = B_1B_2$ i $T = BC$. Przeanalizuj znaczenie poniższych zapytań i postaraj się znaleźć naturalną interpretację dla relacji i zapytań w języku polskim. Zastanów się, czy są to formuły niezależne od dziedziny. Zapisz równoważne im formuły w algebrze relacji zawsze jeśli to możliwe.

1. $\{a \ : \ (\exists\ b)\ (R(a, b) \ \wedge \neg((\exists a')\ a' > a \ \wedge (\exists b')\ R(a', b')))\}$
2. $\{a, b \ : \ (\forall\ c)\ (T(c, a) \ \vee \ T(c, b) \ \vee \ (\forall d)\ \neg T(c, d))\}$

Rozwiązanie.

1. $\{a \ : \ (\exists\ b)\ (R(a, b) \ \wedge \neg((\exists a')\ a' > a \ \wedge (\exists b')\ R(a', b')))\}$

Wszystkie te elementy $a \in A$, dla których nie istnieje inny element (a', b') dla którego $a' > a$. Czyli zwraca to największy element kolumny A.

To chyba nie jest niezależne od dziedziny? Bo jeśli

R=A, B	
A=Stopień	B=Przedmiot
3	B.D.
2	AnalMat

To dziedzina aktywna ma $B = \{(3, \text{"B.D."}), (2, \text{"AnalMat"})\}$ i możemy wziąć D_1 która ma dziedzinę aktywną z dodatkiem $B \cup \{(5, \text{"Euler"}\}$ a jako D_2 wziąć $B \cup \{(4, \text{"RP1R"}\}$ i wtedy wynik jest różny?

Napisanie tego wyżej w języku algebry relacji to

```
pi R.A - (
   pi R.A
   (R join R.A < R.a
      (rho R.a <- R.A, R.b <-R.B (R)
   )
)
```

2. $\{a, b \;:\; (\forall\ c)\ (T(c, a)\ \lor\ T(c, b)\ \lor\ (\forall d)\ \neg T(c, d))\}$

To są pary elementów (a, b), gdzie $a, b \in C$ które albo zawsze pojawiają się w drugiej kolumnie, albo jeśli nie pojawiają się dla pewnej krotki w drugiej kolumnie, to ten element nigdy nie jest na pierwszym miejscu?

Zwraca te elementy kolumny C, które pojawiają się w drugiej kolumnie T dla wszystkich elementów z B stojących na pierwszym miejscu kolumny T. I tutaj jeśli weźmiemy sobie $T = \emptyset$, to dla każdego $c \in B$ i dla każdego $d \in C$ mamy $\neq T(c, d)$, czyli jest to prawdą nawet jak wtłoczymy coś nieskończonego.

Zadanie 4.
Zdecyduj, czy poniższe równości zachodzą. Zaprezentuj dowód lub kontrprzykład.

- $R \bowtie S = S \bowtie R$
- $R \bowtie (S \bowtie T) = (R \bowtie S) \bowtie T$

Rozwiązanie.
$R \bowtie S = S \bowtie R$

Jeśli kolejność kolumn w joinie nie ma znaczenia, to stwierdzenie jest prawdziwe. Wystarczy zapisać np. w RRK:

$$R \bowtie S \equiv \{t \;:\; (\exists r \in R)\ (\exists s \in S)\ r.(S \cap R) = s.(S \cap R)\ \land t.R = r\ \land t.S = s\}$$

$$S \bowtie R \equiv \{t \;:\; (\exists s \in S)\ (\exists r \in R)\ s.(S \cap R) = r.(S \cap R)\ \land t.S = s\ \land t.R = r\}$$

i zauważyć, że $\land$ i $\exists$ są przemienne.

$R \bowtie (S \bowtie T) = (R \bowtie S) \bowtie T$

Tutaj chyba też mamy prawdę. Zauważmy, że definicja natural join tłumaczy się jako

$$x \in R \cap (S \bowtie T) \iff [x \in R \cap (S \cup T)\ \land x.S = x.T]$$

przy czym $R \cap (S \cup T) = (R \cap S) \cup (R \cap T)$. Co więcej, jeśli dla $b \in S \bowtie T$ mamy $x.(S \bowtie T) = b$, to istnieją s, t takie, że $b.S = s$, $b.T = t$, $s.(S \cap T) = t.(S \cap T)$ oraz $x.S = s \land x.T = t$.

Zapiszmy

$$
\begin{aligned}
R \bowtie (S \bowtie T) \equiv \{a \ :\ & (\exists\, r \in R)(\exists b \in S \bowtie T) \\
& \quad r.(R \cap (S \bowtie T)) = b.(R \cap (S \bowtie T)) \ \wedge a.R = r \ \wedge a.(S \bowtie T) = b\} \equiv \\
\equiv \{a \ :\ & (\exists r \in R)(\exists s \in S)(\exists t \in T) \\
& \quad r.((R \cap S) \cup (R \cap T)) = s.((R \cap S) \cup (R \cap T)) \wedge \\
& \quad r.((R \cap S) \cup (R \cap T)) = t.((R \cap S) \cup (R \cap T)) \wedge \\
& \quad r.S = r.T \ \wedge\ s.(S \cap T) = t.(S \cap T) \ \wedge a.R = r \ \wedge\ a.S = s \ \wedge a.T = t\} \equiv \\
\equiv \{a \ :\ & (\exists r \in R)(\exists s \in S)(\exists t \in T) \\
& \quad r.(R \cap S) = s.(R \cap S) \ \wedge\ r.(R \cap T) = t.(R \cap T) \wedge \\
& \quad r.S = r.T \ \wedge s.(S \cap T) = t.(S \cap T) \ \wedge a.R = r \ \wedge\ a.S = s \ \wedge\ a.T = t\}
\end{aligned}
$$

i podobnie będzie jak napiszemy $(R \bowtie S) \bowtie T$

Zadanie 5.
Baza danych składa się z relacji, jara jara jara

RRD lub RRK

Rozwiązanie.
1. Podaj dane aktorów (pseudonim, imię, nazwisko, rok urodzenia, narodowość), którzy pojawili się w filmach produkowanych tylko w jednym roku.

Troszkę oszukane podejście do RRK:

$$
\begin{aligned}
A - \{a \in A \ :\ & (\exists\ x \in R)\ (\exists\ y \in R - \{x\})\ x.psuedonim = a.pseudonim = y.pseudonim \wedge \\
& (\exists\ f1 \in F)\ (\exists f2 \in F - \{f_1\})\ f1.rokProd \neq f2.rokProd \wedge \\
& \{f1.idf, f2.idf\} = \{x.idf, y.idf\}\ \}
\end{aligned}
$$

Lub po (potencjalnie błędnym) przemieleniu przez prawa de Morgana:

$$
\begin{aligned}
& \{a \in A \ :\ \neg[(\exists\ x \in R)\ (\exists\ y \in R - \{x\})\ x.psuedonim = a.pseudonim = y.pseudonim \wedge \\
& \qquad (\exists\ f1 \in F)\ (\exists f2 \in F - \{f1\})\ f1.rokProd \neq f2.rokProd \wedge \\
& \qquad \{f1.idf, f2.idf\} = \{x.idf, y.idf\}]\ \} \equiv \\
& \equiv \{a \in A\ :\ (\forall\ x \in R)\ (\forall\ y \in R - \{x\})\ x.psuedonim \neq a.psuedonim \neq y.pseudonim \vee \\
& \qquad (\forall\ f1 \in F)(\forall\ f2 \in F - \{f1\}) f1.rokProd = f2.rokProd\ \vee\ \{f1.idf, f2.idf\} \neq \{x.idf, y.idf\}\}
\end{aligned}
$$

2. Podaj pełne krotki filmów, które są najnowszymi filmami reżyserów.

$$
\begin{aligned}
& \{f \in F \ :\ \neg(\exists x \in F)\ f.rezyser = x.rezyser \ \wedge f.rokProd < x.rokProd\} \equiv \\
& \equiv \{f \in F\ :(\forall\ x \in F)\ f.rezyser \neq x.rezyser \ \vee\ f.rokProd >= x.rokProd\}
\end{aligned}
$$

3. Dla każdego filmu znajdź aktora, który dostał najwyższą gażę w tym filmie (został najlepiej opłacony z obsady filmu). W relacji wynikowej podaj pseudonim aktora, idf oraz gażę.

$\{ps, i, g \;:\; (\exists\ rl)\ R(ps, i, rl, g) \;\wedge\; (\forall\ ps2, rl2, g2)\ \neg R(ps2, i, rl2, g2) \;\vee\; g2 <= g\}$

4. Podaj pełne krotki aktorów, którzy nigdy nie obniżyli swojej minimalnej gaży (w późniejszych latach mogła ona najwyżej rosnąć). Na wynik nie wpływają lata, w których aktor nie podał minimalnej gaży.

$$\{a \in A \;:\; \neg(\exists\ x, y \in R)\ x.pseudonim = a.pseudonim = y.pseudonim \;\wedge\; x.gaza > y.gaza \;\wedge\; (\exists f, g \in F)\ f.idf = x.idf \;\wedge\; g.idf = y.idf \;\wedge\; f.rokProd < g.rokProd\} \equiv$$