

Zapytania relacyjnego rachunku dziedzin są wyrażeniami postaci

$$\{x_1, \dots, x_n \mid \varphi(x_1, \dots, x_n)\}$$

gdzie  $\varphi$  jest formułą logiki I rzędu ze zmiennymi wolnymi  $x_1, \dots, x_n$ .

- 1. (2 pkt.)** 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. Zapisz równoważne im formuły w algebrze relacji. Rozstrzygnij dla każdego z zapytań czy może się zdarzyć, że jego wynik (tj. zbiór krotek spełniających zapytanie) jest nieskończony.

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

- 2. (3 pkt. - po 0.5 pkt. za podpunkt)** Baza danych składa się z relacji:

- $F(idf, tytuł, reżyser, rokProd, czas)$  —  $idf$  jest kluczem; tytuł i inne atrybuty nie muszą być unikalne;  $czas$  oznacza czas trwania filmu i jest podany w minutach;
- $S(idf, sala, data, godz)$  — w podanej sali i terminie jest projekcja filmu o podanym identyfikatorze;
- $A(pseudo, imię, nazwisko, narodowość, rokUr)$  — informacje o aktorach; pseudonim jest unikalny;
- $R(pseudo, idf, postać, gaza)$  — informacja, że aktor o podanym pseudonimie grał w filmie daną postać i otrzymał za to podaną gażę.
- $M(pseudo, rok, minGaza)$  — informacja, że aktor o podanym pseudonimie w danym roku na podanym poziomie ustalił minimalną gażę za grę w filmie.

Zapisz poniższe zapytania w relacyjnym rachunku dziedzin.

1. Podaj dane aktorów (pseudonim, imię, nazwisko, rok urodzenia, narodowość), którzy pojawili się w filmach produkowanych tylko w jednym roku (powiedzmy, że są to gwiazdy jednego sezonu).
2. Podaj pełne krotki filmów, które są najnowszymi filmami reżyserów.

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.
4. Podaj sale, w których odbyła się projekcja każdego filmu reżysera "Olańskiego". Załóż, że w każdej sali jest jakiś seans.
5. Podaj pełne krotki aktorów, którzy nigdy nie obniżyli swojej minimalnej gaży (w późniejszych latach mogła ona najwyżej rosnać). Na wynik nie wpływają lata, w których aktor nie podał minimalnej gaży.
6. Podaj tytuły filmów, w których zagrał ktoś, kto nie grał w filmie "Roll".

**3.(3pkt.)** <sup>1</sup> Przyjmijmy taką interpretacją wartości NULL, w której oznacza ona *jakaś wartość odpowiedniego typu*, tzn. wiemy, że taka wartość istnieje ale nie wiemy jaka ona jest. Przy takim założeniu wygodne jest zapisywanie NULLi za pomocą zmiennych tzn. jeśli w relacji o atrybutach (`Imię:String`, `Zarobki:Int`) jest krotka (`Józek`, `x`) to oznacza to, że `Józek` ma jakieś zarobki, które można wyrazić pewną wartością typu `Int`, ale nie wiemy jaką. Zakładamy, że każda zmienna może wystąpić w bazie danych co najwyżej jeden raz.

Niech  $D$  będzie relacją ze zmiennymi. Oznaczmy przez  $\text{rep}(D)$  następujący zbiór relacji

$$\{v(D) \mid v \text{ jest wartościowaniem wszystkich zmiennych z } D\}$$

O  $\text{rep}(D)$  należy myśleć, że jest zbiorem wszystkich *zupelných* relacji (tj. relacji bez zmiennych) reprezentowanych przez  $D$ . Na przykład, jeśli  $D$  zawiera wyłącznie krotkę (`Józek`, `x`) to  $\text{rep}(D)$  zawiera wszystkie relacje z dokładnie jedną krotką postaci (`Józek`, `n`), gdzie  $x$  została zwartościowana liczbą całkowitą  $n \in \text{Int}$ .

Oczywiście, żeby ta cała zabawa z NULLami miała sens możemy używać wyłącznie takich wyrażeń algebry relacji  $Q$ , że dla dowolnej relacji  $D$  istnieje relacja (ze zmiennymi)  $Q_D$  reprezentująca wynik  $Q$  na  $D$  tzn. taka, że  $\text{rep}(Q_D) = Q(\text{rep}(D))$ , przy czym przez  $Q(\text{rep}(D))$  oznaczamy obraz zbioru  $\text{rep}(D)$  przez  $Q$ .

Pokaż przykład relacji  $D$  i przykład zapytania  $Q$  będącego pojedynczą selekcją taką, że nie istnieje reprezentacja wyniku  $Q$  na  $D$  tj. nie istnieje relacja (ze zmiennymi)  $Q_D$ , taka że  $\text{rep}(Q_D) = Q(\text{rep}(D))$ . Oznacza to, że w tym systemie nie można używać zapytań z selekcją. W podobny sposób można pokazać, że nie da się też reprezentować wyników złączenia naturalnego. *Wskazówka: zauważ, że (nie)pustość relacji nie zależy od wartościowania jej zmiennych.*

---

<sup>1</sup>rozwiązanie jest dość proste tylko trzeba zrozumieć treść

4.(2pkt.) Baza składa się z relacji:

- $Ludzie(id, nazwisko, rok)$ — $id$  jest kluczem,  $rok$  jest rokiem urodzenia (im większy tym osoba młodsza);
- $M(matka, dziecko)$  — matka jest matką osoby dziecko, zarówno matka, jak i dziecko są kluczami obcymi wskazującymi na relację  $Ludzie$ ;
- $T(ojciec, dziecko)$  — ojciec jest ojcem osoby dziecko, zarówno ojciec, jak i dziecko są kluczami obcymi wskazującymi na relację  $Ludzie$ .

W bazie przechowujemy dane o wszystkich osobach od stworzenia świata (z Adamem i Ewą włącznie).

1. (a) Ile elementów ma zbiór (zakładając poprawność bazy danych)?

$$\{x \mid \exists z, t \ T(z, x) \wedge T(t, z) \wedge M(x, t)\}$$

- (b) Napisz zapytanie relacyjnego rachunku dziedzin wybierające tych ludzi, którzy są starsi od całego swojego rodzeństwa (w tym przyrodniego).
  - (c) Napisz zapytanie relacyjnego rachunku dziedzin wybierające takie matki, które zostały matką wcześniej niż ich matka.
2. Napisz zapytanie relacyjnego rachunku dziedzin wybierające tych ludzi, którzy mają wspólne rodzeństwo (być może przyrodnie) z dowolną osobą o nazwisku 'Kowalski'. Przyjmujemy tu, że każdy jest swoim własnym rodzeństwem.