

Bazy Danych 2024

Lista zadań nr 2

- 1 (1 pkt) 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 .
- 2 (1 pkt) Rozważmy relacje $R(A, B, C)$ 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 .
- 3 (1 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. Zastanów się, czy są to formuły niezależne od dziedziny. Zapisz równoważne im formuły w algebrze relacji zawsze jeśli to możliwe.
 1. $\{a \mid (\exists b)(R(a, b) \wedge \neg((\exists a')a' > a \wedge (\exists b')(R(a', b'))))\}$
 2. $\{a, b \mid (\forall c)(T(c, a) \vee T(c, b) \vee (\forall d)(\neg T(c, d)))\}$
- 4 (1 pkt) Zdecyduj czy poniższe równości zachodzą. Zaprezentuj dowód lub kontrprzykład.
 - $R \bowtie S = S \bowtie R$
 - $R \bowtie (S \bowtie T) = (R \bowtie S) \bowtie T$
- 5 (1 pkt) 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 rrd lub rrk.

1. Podaj dane aktorów (pseudonim, imię, nazwisko, rok urodzenia, narodowość), którzy pojawili się w filmach produkowanych tylko w jednym roku.
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 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 (1 pkt.) W tym zadaniu w krotkach relacji oprócz normalnych stałych takich jak 6 czy "Bazy danych" możemy używać zmiennych. Przyjmijmy taką interpretację, że jeśli w krotce znajduje się zmienna to znaczy, że jest tam *jakaś wartość odpowiedniego typu*, tzn. wiemy, że taka wartość istnieje ale nie wiemy jaka ona jest.

Na przykład - 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 (co jak się okazuje dość dobrze przybliża ideę wartości NULL z SQLa).

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 chcielibyśmy móc 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.

człowiek

Maksymilian: zauważ, że (nie)pustość relacji nie zależy od wartościowania