

Bazy Danych 2019

Lista zadań nr 1

W poniższych zadaniach zakładamy, że relacje nie mogą zawierać wartości NULL, chyba że treść zadania stanowi inaczej.

Z1. (2 pkt.) Czy operator różnicy \setminus da się wyrazić za pomocą wyrażeń algebry relacji z operatorami π , σ , ρ , \times , \cup ? Załóż, że warunki F są formułami zbudowanymi przy użyciu koniunkcji, alternatywy i negacji oraz zawierają wyłącznie atomy postaci $Atr_1 = \text{const}$ lub $Atr_1 = Atr_2$, gdzie Atr_1, Atr_2 są atrybutami, a const stałą odpowiedniego typu. Wskazówka: poszukaj pewnej charakterystycznej cechy, którą mają wszystkie zapytania wyrażalne za pomocą π , σ , ρ , \times , \cup , a której nie musi mieć zapytanie wyrażone z użyciem \setminus .

Z2. (1 pkt.) X, Y i Z są relacjami zawierającymi pojedynczą kolumnę o nazwie A . Student ma napisać wyrażenie algebry relacji wyliczające wartość $X \cap (Y \cup Z)$ nie używając operatorów sumy i przekroju relacji. W bazie danych rozwiązań zadań z poprzednich edycji kursu znalazł następujące wyrażenie:

$$\pi_{X.A}(\sigma_{X.A=Y.A \vee X.A=Z.A} (X \times \rho_{Y(A_Y)}(Y) \times \rho_{Z(A_Z)}(Z)))$$

Czy powinien użyć tego rozwiązania? Jeśli zapytanie jest poprawne to uzasadnij to, jeśli nie to zastanów się czy można je poprawić.

Z3. (2 pkt.) Niech R będzie relacją o atrybutach $X_1, \dots, X_n, Y_1, \dots, Y_m$, a $A \in \{Y_1, \dots, Y_m\}$. Niech f_A będzie funkcją *agregującą*, która dla dowolnego podzbioru krotek relacji R zgadzających się na atrybutach X_1, \dots, X_n agreguje wartości atrybutu A , np. funkcja sum_A zwraca sumę wszystkich wartości na atrybucie A , count_A zlicza ile ich jest, a min_A zwraca wartość minimalną.

W poniższej definicji operatora agregacji i grupowania G używamy standardowych operatorów algebry relacji.

$$G_{f_A(X_1, \dots, X_n)}(R) = \{(x_1, \dots, x_n, a) \mid (x_1, \dots, x_n) \in \pi_{X_1, \dots, X_n}(R) \text{ oraz} \\ a = f_A(\sigma_{X_1=x_1 \wedge \dots \wedge X_n=x_n}(R))\}$$

Algebrę relacji rozszerzoną o operator G oznaczamy RA_G . Będziemy używać bazy danych składającej się z relacji B, P, L takich, że $B(\text{osoba}, \text{bar})$, oznacza, że *osoba* bywa w *bar*, $P(\text{sok}, \text{bar}, \text{cena})$ oznacza, że w *bar* podają *sok* w cenie *cena* za porcję, a $L(\text{osoba}, \text{sok})$ oznacza, że *osoba* lubi *sok*.

Przykład Zapytanie $G_{\text{count}_{\text{osoba}(\text{bar})}}(B)$ zwraca relację zawierającą pary (bar, n) , takie, że w barze *bar* bywa dokładnie n osób.

a) (0.5pkt) Co zwraca zapytanie $G_{\text{count}_{\text{sok}(\text{bar})}}(\pi_{\text{bar}, \text{sok}}(B \bowtie P))$?

b) (0.5pkt) Napisz zapytanie RA_G zwracające pary $(\text{osoba}, \text{bar})$, takie, że dana osoba lubi co najmniej 5 soków podawanych w danym barze.

- c) (0.5pkt) Napisz zapytanie RA_G zwracające krotki (*osoba, sok, cena*) takie, że w barach, w których bywa osoba *osoba* sok *sok* można najtaniej kupić za *cena*.
- d) (0.5pkt) Napisz zapytanie RA_G zwracające krotki (*osoba, sok, bar*) takie, że wśród wszystkich barów, w których bywa osoba *osoba* sok *sok* jest najtańszy w barze *bar*.

W poniższych dwóch zadaniach będziemy korzystać z bazy składającej się z następujących relacji (klucze główne są podkreślone):

- $K = (\underline{nrK}, \text{autor}, \text{tytuł}, \text{opis})$ — książki, jako dzieła, a nie konkretne egzemplarze;
- $E = (\underline{sygn}, \underline{nrK}, \text{rokWyd})$ — konkretne egzemplarze książek z rokiem wydania i odnośnikiem do tabeli K, gdzie po atrybucie nrK można odszukać dzieło, znaleźć jego autora itp.;
- $C = (\underline{nrC}, \text{nazwisko}, \text{adres})$ — dane czytelników zarejestrowanych w bibliotece; dla uproszczenia przyjmujemy, że adres to tylko nazwa miasta;
- $W = (\underline{nrC}, \underline{sygn}, \text{czasWyp}, \text{czasZwr})$ — informacja o wypożyczeniu egzemplarza książki; pierwsze pole *czas* to moment wypożyczenia; drugie pole *czas* jest na początku puste i wypełnia się je dopiero w chwili zwrotu książki; czasy zapisywane są z bardzo dużą dokładnością i przyjmujemy, że nie jest możliwe wykonanie dwóch akcji w bazie w tym samym czasie;
- $L = (\underline{nrC}, \underline{nrK})$ — wpis w tej tabeli oznacza, że dany czytelnik lubi dane dzieło.

Z4. (2 pkt.) Potrenuj samodzielnie pisząc wyrażenia algebry relacji dla poniższych pytań. Na zajęciach trzeba będzie zapisać w algebrze relacji podobne, ale nie takie same zapytania.

- Podaj nazwiska osób, które choć raz wypożyczyły pewną książkę (dzieło) tego samego dnia, gdy ją oddały (to samo dzieło). Zastanów się, jak znaleźć czytelników, którzy od chwili zapisania do biblioteki cały czas mają wypożyczoną jakąś książkę. Aby z czasu wypożyczenia odzyskać dzień, użyj funkcji $day(\text{czasWyp})$.
- Podaj nazwiska osób, które tylko raz wypożyczyły książkę (i oddały ją).
- Podaj nazwiska czytelników, którzy wypożyczyli jakąś książkę, której nie lubią.
- Podaj tytuły książek lubianych przez te same osoby, które lubią *Lalkę* Bolesława Prusa.