

Bazy Danych 2017

Lista zadań nr 1

7.03.2017

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 koniunkcjami atomów postaci $\text{Attr}_1 = \text{const}$ lub $\text{Attr}_1 = \text{Attr}_2$, gdzie $\text{Attr}_1, \text{Attr}_2$ 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 . Zastanów się też czy ograniczenie na postać warunków F powyżej jest konieczne? Czy dopuszczenie alternatywy lub nierówności (\neq) w warunkach F coś zmienia?

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:

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

Czy powinien użyć tego rozwiązania? A co jeśli dopuścimy wartości NULL?

Z3. (2 pkt.) Niech R będzie relacją o atrybutach $X_1, \dots, X_n, Y_1, \dots, Y_m$, a f_A będzie funkcją agregującą, która dla każdego podzbioru krotek relacji R zgadzających się na atrybutach X_1, \dots, X_n zwraca wartość typu A . 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 $B(\textit{osoba}, \textit{bar})$, oznacza, że *osoba* bywa w *bar*, $P(\textit{sok}, \textit{bar}, \textit{cena})$ oznacza, że w *bar* podają *sok* w cenie *cena* za porcję, a $L(\textit{osoba}, \textit{sok})$ oznacza, że *osoba* lubi *sok*.

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

- a) (0.5pkt) Co zwraca zapytanie $G_{\text{count}_{\text{int}}(\textit{bar})}(\pi_{\textit{bar}, \textit{sok}}(B \bowtie P))$?
- b) (0.5pkt) Napisz zapytanie RA_G zwracające pary $(\textit{osoba}, \textit{bar})$, takie, że dana osoba lubi co najmniej 5 soków podawanych w danym barze.
- c) (0.5pkt) Napisz zapytanie RA_G zwracające listę soków, z których każdy jest podawany przynajmniej w 5 barach, a przy tym lubi go przynajmniej 5 osób.

- 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):

- $K = (\underline{nrK}, autor, tytuł, opis)$ — książki, jako dzieła, a nie konkretne egzemplarze;
- $E = (\underline{sygn}, nrK, rokWyd)$ — konkretne egzemplarze książek z rokiem wydania i odnośnikiem do tabeli K, gdzie po atrybucie nrK można odszukać dzieło, znaleźć jego autora itp.;
- $C = (\underline{nrC}, nazwisko, adres)$ — dane czytelników zarejestrowanych w bibliotece; dla uproszczenia przyjmujemy, że adres to tylko nazwa miasta;
- $W = (\underline{nrC}, sygn, \underline{czasWyp}, \underline{czasZwr})$ — informacja o wypożyczeniu egzemplarza książki; pierwsze pole $czas$ to moment wypożyczenia; drugie pole $czas$ jest na początku puste i wypełnia się 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.) Wyjaśnij po polsku znaczenie zapytań:

- (a) $\pi_{autor}((K \bowtie E) \setminus \pi_{nrK, autor, tytuł, opis, sygn, rokWyd}(\sigma_{czasWyp > 31.12.2006}(K \bowtie E \bowtie W)))$;
- (b) $\pi_{nazwisko}(\sigma_{sygn1 \neq sygn}(\sigma_{czasZwr \text{ IS NULL and } czasZwr1 \text{ IS NULL}}(W \bowtie C \bowtie \rho_{W1(nrC, sygn1, czasWyp1, czasZwr1)}(W))))$.
- (c) $\pi_{nrK, sygn}(E) \setminus \pi_{nrK, sygn}(\sigma_{r1 < r2}(\rho_{E1(s1, nrK, r1)}(E) \bowtie \rho_{E1(sygn, nrK, r2)}(E)))$

Z5. (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.

- (a) 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 $day(czasWyp)$.
- (b) Podaj nazwiska osób, które tylko raz wypożyczyły książkę (i oddały ją).
- (c) Podaj nazwiska czytelników, którzy wypożyczyli jakąś książkę, której nie lubią.
- (d) Podaj tytuły książek lubianych przez te same osoby, które lubią *Lalkę* Bolesława Prusa.