Lista 3

13 March, 2024 22:45

zad.	1	2	3	4	5	6	7	suma
pkt.								
maxpkt.	1	1	1	1	1	1	0,5	

Lista zadań nr 3

Bazy Danych 2024

Zapytania koniunkcyjne to zapytania rrd zbudowane z formuł atomowych (np. $R(x, y)$) oraz koniunkcji i kwantyfikatorów egzystencjalnych. W ogólności formułami atomowymi mogą być również równości i nierówności między stałymi i zmiennymi (np. $x = 5$, $x \neq y$, $x < z$) ale na tej liście pozwalamy wyłącznie na atomy relacyjne oraz nie pozwalamy na używanie stałych.

Rozważmy bazę danych reprezentującą pewien graf skierowany o krawędziach zapisanych w relacji $E(S, T)$. Niestety w naszych zapytaniach nie możemy używać relacji E. W zamian mamy dostęp do relacji $P_i(x, y)$ dla pewnych $i > 1$. Relacja $P_i(x, y)$ zawiera pary wierzchołków połączone ścieżką długości i, np. $P_2(x, z)$ mogłaby by zdefiniowana jako $(\exists y)E(x, y) \wedge E(y, z)$. Odpowiada to sytuacji, w której np. ze względów bezpieczeństwa dostęp do bazy danych mamy wyłącznie za pomocą zestawu perspektyw (widoków), a dostęp do oryginalnych relacji jest zablokowany.

Jeśli chcemy wyliczyć odpowiedzi na jakieś zapytanie ψ używające relacji E możemy spróbować zmodyfikować (przepisać) ψ tak aby zamiast E wykorzystać symbole dostępnych perspektyw. Np. jeśli mamy wyłącznie dostęp do perspektywy $P_2(x, y)$, a chcemy zapisać zapytanie $P_4(x, z) = (\exists y_1, y_2, y_3)E(x, y_1) \wedge E(y_1, y_2) \wedge E(y_2, y_3) \wedge E(y_3, z)$ możemy to zrobić tak: $P_4'(x, z) = (\exists y)P_2(x, y) \wedge P_2(y, z)$ (zauważ, że P_4' też jest zapytaniem koniunkcyjnym i jest równoważne $P_4(x, y)$).

Na rozgrzewkę pokaż, jak przepisać zapytanie $P_7(x, y)$ używając wyłącznie perspektyw $P_2(x, y)$ i $P_3(x, y)$.

1. (1 pkt.) Pokaż, że nie istnieje takie zapytanie koniunkcyjne używające jako formuł atomowych wyłącznie perspektyw $P_3(x, y)$ i $P_4(x, y)$, które jest równoważne zapytaniu $P_5(x, y)$.

2. (1 pkt.) Napisz zapytanie rrd (dozwolone $\exists, \forall$ i wszystkie spójniki boolowskie), które korzysta wyłącznie z perspektyw $P_3(x, y)$ i $P_4(x, y)$ i jest równoważne zapytaniu $P_5(x, y)$.

Dziedziną bazy danych A, $\mathrm{dom}(A)$, nazwiemy zbiór wszystkich elementów zawartych w krotkach w relacjach z A. Homomorfizm pomiędzy bazami danych A i B to funkcja $h : \mathrm{dom}(A) \to \mathrm{dom}(B)$ spełniająca warunek: dla każdego symbolu relacji R i każdej krotki $(a_1, \ldots, a_n) \in \mathrm{dom}(A)^n$ (gdzie n jest liczbą atrybutów R) zachodzi

$$A \models R(a_1, \ldots, a_n) \implies B \models R(h(a_1), \ldots, h(a_n))$$

3. (1 pkt) W czasie wykładu pojawił się dowód następującej implikacji.

Niech Q_1 i Q_2 będą boolowskimi zapytaniami koniunkcyjnymi (tzn. bez zmiennych wolnych). Jeśli istnieje homomorfizm ze bazy kanonicznej C_{Q_2} w C_{Q_1} to $Q_1 \subseteq Q_2$.

Udowodnij implikację w drugą stronę. Możesz skorzystać z faktu, że dla dowolnego boolowskiego zapytania koniunkcyjnego Q i bazy danych D mamy $D \models Q$ (tzn. zapytanie jest Q spełnione w D) wtedy i tylko wtedy gdy istnieje homomorfizm ze bazy kanonicznej C_Q w bazę D.

4. (1 pkt)
1. Pokaż, że istnieje graf G, taki, że dla dowolnego grafu G' istnieje homomorfizm z G' w G.
2. Udowodnij lub podaj kontrprzykład na stwierdzenie: dla dowolnych grafów G_1 i G_2 następujące warunki są równoważne:
   - istnieją homomorfizmy $h_1 : G_1 \to G_2$ oraz $h_2 : G_2 \to G_1$ (mówimy, że takie grafy są *homomorficznie równoważne*)
   - istnieje izomorfizm $f : G_1 \to G_2$.

5. (1 pkt) W tym zadaniu rozważamy grafy symetryczne tzn. takie, że jeśli istnieje krawędź z v do w to istnieje też krawędź z w do v. Pokaż, że każde dwa takie grafy będące cyklami o parzystej długości są homomorficznie równoważne.

6. (1 pkt.) Mówimy, że formuła rachunku zdań jest w postaci 3CNF, gdy jest w koniunkcyjnej postaci normalnej, a dodatkowo każda klazula zawiera najwyżej 3 zmienne. Przykład formuły 3CNF: $(p \vee \neg q \vee r) \wedge (s \vee \neg r)$.

Rozważmy problem spełnialności formuł rachunku zdań zwany 3SAT: dla danej formuły w postaci 3CNF, rozstrzygnij czy istnieje wartościowanie ją spełniające.

Pokaż, że problem 3SAT można wyrazić jako problem istnienia homomorfizmu pomiędzy bazami danych, tzn. pokaż, że jeśli potrafisz rozwiązywać problem istnienia homomorfizmu to potrafisz także rozwiązywać problem 3SAT.

Twoja konstrukcja powinna działać w czasie wielomianowym.

7. (0.5 pkt.) Wiadomo, że problem 3SAT jest trudny obliczeniowo. Za wykazanie, że istnieje jakikolwiek wielomianowy algorytm rozwiązujący ten problem Clay Institute ufundował nagrodę w wysokości miliona dolarów. Problem ewaluacji zapytań koniunkcyjnych (czyli istnienia homomorfizmu) nie jest uznawany za szczególnie trudny, a silniki bazodanowe od wielu lat wydajnie wyliczają odpowiedzi na zapytania SQL. Tymczasem w zadaniu 6 tak naprawdę pokazaliśmy, że problem 3SAT jest łatwiejszy (lub równie trudny) jak problem istnienia homomorfizmu. Wyjaśnij dlaczego to nie jest paradoks (a rozwiązanie zadania 6 nie wystarczy aby otrzymać nagrodę).

1,2

13 March, 2024 22:47

Lista zadań nr 3

Bazy Danych 2024

Zapytania koniunkcyjne to zapytania rrd zbudowane z formuł atomowych (np. $R(x, y)$) oraz koniunkcji i kwantyfikatorów egzystencjalnych. W ogólności formułami atomowymi mogą być również równości i nierówności między stałymi i zmiennymi (np. $x = 5$, $x \neq y$, $x < z$) ale na tej liście pozwalamy wyłącznie na atomy relacyjne oraz nie pozwalamy na używanie stałych.

Rozważmy bazę danych reprezentującą pewien graf skierowany o krawędziach zapisanych w relacji $E(S, T)$. Niestety w naszych zapytaniach nie możemy używać relacji E. W zamian mamy dostęp do relacji $P_i(x, y)$ dla pewnych $i > 1$. Relacja $P_i(x, y)$ zawiera pary wierzchołków połączone ścieżką długości i, np. $P_2(x, z)$ mogłaby by zdefiniowana jako $(\exists y)E(x, y) \wedge E(y, z)$. Odpowiada to sytuacji, w której np. ze względów bezpieczeństwa dostęp do bazy danych mamy wyłącznie za pomocą zestawu perspektyw (widoków), a dostęp do oryginalnych relacji jest zablokowany.

Jeśli chcemy wyliczyć odpowiedzi na jakieś zapytanie ψ używające relacji E możemy spróbować zmodyfikować (przepisać) ψ tak aby zamiast E wykorzystać symbole dostępnych perspektyw. Np. jeśli mamy wyłącznie dostęp do perspektywy $P_2(x, y)$, a chcemy zapisać zapytanie $P_4(x, z) = (\exists y_1, y_2, y_3)E(x, y_1) \wedge E(y_1, y_2) \wedge E(y_2, y_3) \wedge E(y_3, z)$ możemy to zrobić tak: $P_4'(x, z) = (\exists y)P_2(x, y) \wedge P_2(y, z)$ (zauważ, że P_4' też jest zapytaniem koniunkcyjnym i jest równoważne $P_4(x, y)$).

Na rozgrzewkę pokaż, jak przepisać zapytanie $P_7(x, y)$ używając wyłącznie perspektyw $P_2(x, y)$ i $P_3(x, y)$.

1. (1 pkt.) Pokaż, że nie istnieje takie zapytanie koniunkcyjne używające jako formuł atomowych wyłącznie perspektyw $P_3(x, y)$ i $P_4(x, y)$, które jest równoważne zapytaniu $P_5(x, y)$.

2. (1 pkt.) Napisz zapytanie rrd (dozwolone $\exists, \forall$ i wszystkie spójniki boolowskie), które korzysta wyłącznie z perspektyw $P_3(x, y)$ i $P_4(x, y)$ i jest równoważne zapytaniu $P_5(x, y)$.

Dziedziną bazy danych A, $\mathrm{dom}(A)$, nazwiemy zbiór wszystkich elementów zawartych w krotkach w relacjach z A. Homomorfizm pomiędzy bazami danych A i B to funkcja $h : \mathrm{dom}(A) \to \mathrm{dom}(B)$ spełniająca warunek: dla każdego symbolu relacji R i każdej krotki $(a_1, \ldots, a_n) \in \mathrm{dom}(A)^n$ (gdzie n jest liczbą atrybutów R) zachodzi

$$A \models R(a_1, \ldots, a_n) \implies B \models R(h(a_1), \ldots, h(a_n))$$

3. (1 pkt) W czasie wykładu pojawił się dowód następującej implikacji.

Niech Q_1 i Q_2 będą boolowskimi zapytaniami koniunkcyjnymi (tzn. bez zmiennych wolnych). Jeśli istnieje homomorfizm ze bazy kanonicznej C_{Q_2} w C_{Q_1} to $Q_1 \subseteq Q_2$.

Udowodnij implikację w drugą stronę. Możesz skorzystać z faktu, że dla dowolnego boolowskiego zapytania koniunkcyjnego Q i bazy danych D mamy $D \models Q$ (tzn. zapytanie jest Q spełnione w D) wtedy i tylko wtedy gdy istnieje homomorfizm ze bazy kanonicznej C_Q w bazę D.

4

13 March, 2024 22:47

4. (1 pkt) 1. Pokaż, że istnieje graf G, taki, że dla dowolnego grafu G' istnieje homomorfizm z G' w G.

2. Udowodnij lub podaj kontrprzykład na stwierdzenie: dla dowolnych grafów G_1 i G_2 następujące warunki są równoważne:
   - istnieją homomorfizmy $h_1 : G_1 \to G_2$ oraz $h_2 : G_2 \to G_1$ (mówimy, że takie grafy są *homomorficznie równoważne*)
   - istnieje izomorfizm $f : G_1 \to G_2$.

5

13 March, 2024 22:47

5. (1 pkt) W tym zadaniu rozważamy grafy symetryczne tzn. takie, że jeśli istnieje krawędź z v do w to istnieje też krawędź z w do v. Pokaż, że każde dwa takie grafy będące cyklami o parzystej długości są homomorficznie równoważne.

6,7

13 March, 2024 22:47

6. (1 pkt.) Mówimy, że formuła rachunku zdań jest w postaci 3CNF, gdy jest w koniunkcyjnej postaci normalnej, a dodatkowo każda klazula zawiera najwyżej 3 zmienne. Przykład formuły 3CNF: $(p \vee \neg q \vee r) \wedge (s \vee \neg r)$.

Rozważmy problem spełnialności formuł rachunku zdań zwany 3SAT: dla danej formuły w postaci 3CNF, rozstrzygnij czy istnieje wartościowanie ją spełniające.

Pokaż, że problem 3SAT można wyrazić jako problem istnienia homomorfizmu pomiędzy bazami danych, tzn. pokaż, że jesli potrafisz rozwiązywać problem istnienia homomorfizmu to potrafisz także rozwiązywać problem 3SAT.

Twoja konstrukcja powinna działać w czasie wielomianowym.

7. (0.5 pkt.) Wiadomo, że problem 3SAT jest trudny obliczeniowo. Za wykazanie, że istnieje jakikolwiek wielomianowy algorytm rozwiązujący ten problem Clay Institute ufundował nagrodę w wysokości miliona dolarów. Problem ewaluacji zapytań koniunkcyjnych (czyli istnienia homomorfizmu) nie jest uznawany za szczególnie trudny, a silniki bazodanowe od wielu lat wydajnie wyliczają odpowiedzi na zapytania SQL. Tymczasem w zadaniu 6 tak naprawdę pokazaliśmy, że problem 3SAT jest łatwiejszy (lub równie trudny) jak problem istnienia homomorfizmu. Wyjaśnij dlaczego to nie jest paradoks (a rozwiązanie zadania 6 nie wystarczy aby otrzymać nagrodę).