**Lista zadań nr 4** 30.03.2021

Rozważamy bazę danych zawierającą relację E(X, Y) opisującą krawędzie pewnego grafu skierowanego.

**1. (1 pkt.)** Dla danego grafu $G$ skonstruuj takie zapytanie koniunkcyjne $Q()$ oraz bazę danych $D$, że $Q()$ jest prawdziwe w bazie $D$ wtw gdy graf $G$ jest 3-kolorowalny. Dlaczego, choć wiadomo, że problem istnienia 3-kolorowania grafu jest trudny (milion dolarów za efektywny algorytm!) to problem ewaluacji zapytań koniunkcyjnych (a nawet zapytań SQL) nie jest uznawany za szczególnie trudny?

**2. (1 pkt.)** Rozważmy następujące zapytanie w Datalogu.

```
T(X, Y) :- E(X,Y).
T(X, Y) :- T(X,Z), T(Z,Y).
```

Przypomnij definicję semantyki dla Datalogu, a następnie pokaż, że dla każdego $i \in \mathbb{N}_+$ zachodzi $T^i = \{(a, b) |$ istnieje ścieżka z $a$ do $b$ o długości $\leq 2^i\}$

**3. (2 pkt., po 0.5 za podpunkt)** Napisz następujące zapytania datalogowe. Użyj stałych $n$ i $m$ tam gdzie jest to potrzebne.

1. Zwróć wierzchołki, do których można dojść scieżką z $n$ lub ścieżką z $m$.
2. Zwróć wierzchołki, do których można dojść ścieżką z $n$ i ścieżką z $m$.
3. Zwróć pary wierzchołków, do których można dojść z wierzchołka $n$ ścieżkami o tej samej długości.
4. Zwróć pary wierzchołków $x, y$, takie, że do $x$ oraz do $y$ można dojść z wierzchołka $n$ ścieżkami, które mają różną długość.

**Definicja.** Graf jest $k$-kolorowalny jeśli każdemu wierzchołkowi tego grafu możemy przyporządkować jeden z $k$ kolorów w taki sposób aby każde dwa wierzchołki połączone krawędzią miały różne kolory.

**4. (1 pkt.)** Wiadomo, że graf jest 2-kolorowalny wtedy i tylko wtedy gdy nie zawiera cyklu o nieparzystej długości. Napisz zapytanie datalogowe $Q()$ spełnione w grafach, które nie są 2-kolorowalne.

**5. (1 pkt.)**
1. Czy można napisać zapytanie datalogowe spełnione wtw gdy w grafie nie ma ścieżki pomiędzy wyróżnionymi wierzchołkami $n$ i $m$?
2. Czy można napisać zapytanie datalogowe spełnione wtw gdy graf zawiera parzystą liczbę wierzchołków?

**6.* (1 pkt., dla chętnych, trudne!)** Rozważmy własność $\mathcal{P}(a, b)$: Czy w grafie istnieje ścieżka (niekoniecznie prosta) pomiędzy $a$ i $b$, której długość jest kwadratem pewnej liczby naturalnej.

1. Zauważ, że istnienie takiej ścieżki jest zachowywane dla rozszerzeń grafu (jesli $G$ spełnia $\mathcal{P}(a, b)$ to dowolny nadgraf $H \supseteq G$ też) oraz dla jego obrazów przez homomorfizm (jesli $G \models \mathcal{P}(a, b)$ to dla dowolnego grafu $H$ t.że istnieje $h : G \to H$ również $H \models \mathcal{P}(h(a), h(b))$). A więc *proste* metody na pokazywanie, że czegoś się nie da wyrazić nie pomogą.

2. Pokaż, że własności $\mathcal{P}(a, b)$ nie da się wyrazić w datalogu. Wskazówka: rozważaj grafy będące prostymi ścieżkami pomiędzy $a$ i $b$ –nazwijmy je *słowami*. Wykorzystaj technikę *pompowania* czyli spostrzeżenia w rodzaju: dla danego programu datalogowego $\pi$ jeśli akceptuje on *słowo* $w$ dłuższe niż pewne $N \in \mathbb{N}$, które zależy wyłącznie od rozmiaru $\pi$ to program $\pi$ akceptuje też inne *napompowane* słowo np. $w_1uw_2$ takie, że $w_1, w_2$ są podsłowami $w$, a $u$ jest dowolnym słowem.