

JEDEN KROK OD MODELU RELACYJNEGO.
KILKA PODSTAWOWYCH ZAGADEK TEORII BAZ DANYCH.
OPIS PROJEKTU BADAWCZEGO (STRESZCZENIE POPULARNONAUKOWE)
JERZY MARCINKOWSKI

Niniejszy projekt mieści się w obszarze teorii baz danych. Teoria ta, inspirowana przez praktykę baz danych, próbuje: • identyfikować, opisywać i badać, w abstrakcyjny matematyczny sposób, podstawowe teoretyczne mechanizmy stojące za praktycznymi aplikacjami w zakresie baz danych; • proponować nowe pojęcia i mechanizmy które mogłyby być pożyteczne dla praktyki baz danych; • tłumaczyć – poprzez dowodzenie wyników negatywnych – porażki niektórych próbowanych w praktyce pomysłów.

W pełnym Opisie Projektu Badawczego szczegółowo omawiamy cztery podobszary/ otwarte techniczne problemy teorii baz danych, którymi planujemy się zajmować. Żaden z tych problemów nie daje się rzetelnie streścić na jednej stronie. Ale spróbujmy trochę opowiedzieć o jednym z nich.

W zwykłej relacyjnej bazie danych gromadzimy pewne znane nam fakty. Na przykład możemy znać następujące fakty dotyczące Zwierząt Plotkujących w Lesie:

Plotkuje(wiewiór, misiek), Plotkuje(wiewiór, wilk), Plotkuje(wilk, lis)

gdzie napis *Plotkuje(wiewiór, wilk)* oznacza że **wiemy** że plotka puszczone przez wiewióra dojdzie w końcu do wilka. Ale uwaga: z naszej przykładowej bazy danych **nie wynika** czy plotka puszczone przez wilka dojdzie do wiewióra – może nie dojdzie, a może dojdzie, tylko my o tym nie wiemy.

We współcześnie proponowanych rozszerzeniach tradycyjnego modelu relacyjnego możemy, oprócz wiedzy o faktach, gromadzić też wiedzę „semantyczną” – o pewnych rządzących światem regułach. Na przykład to co wiemy o plotkowaniu możemy zechcieć sformalizować w dwóch regułach:

$Plotkuje(x, y) \Rightarrow \exists z Plotkuje(y, z) \qquad Plotkuje(x, y), Plotkuje(y, z) \Rightarrow Plotkuje(x, z).$

Pierwsza z powyższych reguł mówi że jeśli do kogoś (do y) w ogóle dochodzą jakieś plotki („istnieje x który plotkuje do y ”), to istnieje również taki z któremu plotkuje ten y . Druga reguła mówi że jeśli do y dochodzą plotki od x zaś do z dochodzą plotki od y to do z dochodzą plotki od x .

Ale w bazie danych wiedzę nie tylko chcemy gromadzić. Chcemy również odpowiadać na zapytania. Na przykład chcemy wiedzieć...

ĆWICZENIE ...czy z naszej wiedzy o faktach i regułach wynika teraz **na pewno** że: (a.) *Plotkuje(wiewiór, lis)*? (b.) *Plotkuje(lis, wiewiór)*? (c.) $\exists x Plotkuje(lis, x)$ (czyli że jest ktoś taki komu lis plotkuje)? (d.) $\exists x Plotkuje(x, wiewiór)$? (Odpowiedzi na dole strony).

Wiele się w ostatnich latach dowiedzieliśmy o algorytmach wyliczających odpowiedzi na takie zapytania – zapytania bazodanowe w obecności reguł. W zależności od dokładnego formatu użytych reguł te algorytmy mogą być obliczeniowo proste, złożone, albo mogą wręcz wcale nie istnieć. Ale jest jedna kwestia którą rozumiemy słabo. Popatrzmy na zapytanie: $\exists x Plotkuje(x, x)$ - tzn. czy istnieje ktoś kogo własne plotki w końcu do niego wracają. Czy to zapytanie jest, zgodnie z naszą zawartą w faktach i regułach wiedzą, na pewno prawdziwe?

Odpowiedź na to pytanie zależy od tego, czy zakładamy (dodatkowo) że w lesie jest tylko skończenie wiele zwierząt (wtedy na pewno jest prawdziwe) czy też nie używamy takiego założenia (wtedy nie mamy co do tego pewności).

Znana nam dziś teoria dotycząca algorytmów odpowiadania na zapytania bazodanowe w obecności reguł zupełnie nie umie sobie radzić z takim dodatkowym, naturalnym przecież, założeniem o skończoności świata. Jednym z celów naszego projektu jest wypełnienie tej luki.

Rozwiązanie ćwiczenia: a. nie; b. tak; c. tak; d. nie.