

# Od silnych logik deskrypcyjnych do fragmentów logiki pierwszego rzędu z wieloma zmiennymi: wnioskowania w strukturach skończonych

Poszukiwanie eleganckich formalizmów o rozstrzygalnym problemie spełnialności i sensownej sile wyrazu to ważny temat w informatyce, o motywacjach płynących z jej rozmaitych obszarów takich, jak weryfikacja programów i sprzętu, sztuczna inteligencja, obliczenia rozproszone czy bazy danych.

Obiekty pojawiające się w zastosowaniach, np. bazy danych/bazy wiedzy, programy komputerowe czy ich obliczenia można modelować za pomocą struktur logicznych, a ich własności opisywać formułami logicznymi. Algorytmy rozwiązujące *problem spełnialności*, tj. rozstrzygające, czy podana na wejściu formuła ma model, mogą być wtedy używane do sprawdzania, czy istnieją obiekty o zadanej specyfikacji, a więc np. do odpowiadania na zapytania do baz danych/baz wiedzy, czy informowania o potencjalnej możliwości niebezpiecznego zachowania programów.

Oczywiście mamy naturalną zależność: im więcej własności można wyspecyfikować w danym formalizmie (im bogatszy język), tym trudniej jest prowadzić automatyczne wnioskowania. Np., naturalny kandydat na język specyfikacji—logika pierwszego rzędu—choć dobrze ugruntowana w matematyce i informatyce i bardzo elegancka, jest zbyt silna w stosunku do rozważanych zastosowań, gdyż już w klasycznych pracach Churcha i Turinga w latach 30-tych XX wieku pokazano, że jej problem spełnialności jest nierozstrzygalny: nie ma (nawet w teorii) algorytmu, który by go rozwiązywał. W tym projekcie planujemy badać potencjalne języki specyfikacji z dwóch grup: ekspresyjne *logiki deskrypcyjne* oraz *fragmenty logiki pierwszego rzędu z wieloma zmiennymi*.

Logiki deskrypcyjne mają dziś niemalże status standardu przemysłowego. Stanowią one np. podstawę OWL (Web Ontology Language) i mają zastosowania w biomedycynie i innych dziedzinach. Skrojone są do zastosowań w Reprezentacji Wiedzy. Podstawowe pojęcia w logikach deskrypcyjnych (DL) to *koncepty* (klasy obiektów) i *role* (relacje pomiędzy obiektami). Konkretne DL są charakteryzowane przez dostępne w nich konstruktory do budowania złożonych konceptów i ról z konceptów i ról atomowych. *Baza wiedzy* składa się z ABoxa, zamierającego fakty dotyczące elementów uniwersum (jest to odpowiednik klasycznej bazy danych) oraz TBoxa, który zawiera wiedzę, z której dodatkowe fakty, nieobecne jawnie w ABoxie mogą być wywnioskowane. Typowy problem decyzyjny w tej dziedzinie to *spełnialność bazy wiedzy*: dla danej bazy sprawdź czy ma ona model (jest niesprzeczna). Inny studiowany problem to *ontology-mediated query answering*, OMQA, w którym na wejściu, oprócz bazy wiedzy dane jest *zapytanie*. Problem polega na sprawdzeniu, czy zapytanie to jest prawdziwe w każdym modelu danej bazy wiedzy. Rozważa się rozmaite klasy zapytań, my będziemy badać głównie klasyczne *zapytania koniunkcyjne* oraz pewne ich generalizacje.

Skoncentrujemy się na logikach deskrypcyjnych o dużej sile wyrazu z rodzin  $\mathcal{Z}$  i  $\mathcal{S}$  i badać będziemy *skończone* wersje spełnialności i OMQA, tj. wersje w których dopuszczamy jedynie modele skończone. To naturalne założenie, gdyż w praktyce obiekty, o których prowadzimy wnioskowania, np. bazy wiedzy są zazwyczaj skończone. O ile problem OMQA dla ważnych logik deskrypcyjnych jest całkiem dobrze zbadany, o tyle jego skończona wersja zaczęła być badana dopiero niedawno. W szczególności dużo jest do zrobienia na temat skończonego OMQA w rodzinie logik  $\mathcal{Z}$  i to od niej zaczniemy nasze badania.

Logiki deskrypcyjne mają swoją własną notację, ale większość z nich może być przetłumaczona na logikę pierwszego rzędu. Oprócz badania logik deskrypcyjnych możemy zatem badać też fragmenty logiki pierwszego rzędu, w których zanurza się takie tłumaczenie. Spełnialność baz wiedzy oraz OMQA przekładają się wtedy na spełnialność formuł w odpowiednich fragmentach. Tłumaczenie to można przeprowadzić używając tylko dwóch zmiennych, my jednak będziemy badać przede wszystkim fragmenty, w których liczba zmiennych jest nieograniczona. Skoncentrujemy się na wprowadzonym ostatnio *triguarded fragment* oraz zdefiniowanej już niemal 50 lat temu, nieco od tego czasu zapomnianej, lecz, jak wierzymy, wartej wznowienia badań, *klasie Mastowa*. Studiowanie fragmentów tego typu pozwala nam lepiej zrozumieć powody rozstrzygalności/nierozstrzygalności w świecie logik deskrypcyjnych, ale jest też interesujące samo w sobie, gdyż niektóre takie fragmenty pozwalają wyrażać własności nieopisywalne w logikach deskrypcyjnych, a nawet mówić o rzeczach, o których te ostatnie mówić nie mogą, jak np. relacje o liczbie argumentów większej od 2. Ostatni argument czyni takie fragmenty bardziej przydatnymi od logik deskrypcyjnych w pewnych zastosowaniach, np. w klasycznej teorii baz danych.