

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

БІЛОІВАНЕНКО МАКСИМ ВІКТОРОВИЧ

УДК 004.8

МЕТОД ТА МОДЕЛІ РОЗШИРЕННЯ СЕМАНТИКИ ВІДНОШЕНЬ ОНТОЛОГІЙ  
ЗА ДОПОМОГОЮ ОПИСУ МАТЕМАТИЧНИХ ЗАЛЕЖНОСТЕЙ МІЖ  
ПАРАМЕТРАМИ ОБ'ЄКТІВ

05.13.23 – системи та засоби штучного інтелекту

Автореферат дисертації на здобуття наукового ступеня кандидата технічних наук

Харків – 2013

Дисертацією є рукопис.

Роботу виконано у Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

Науковий керівник – доктор технічних наук, професор  
Терзіян Ваган Якович,  
Харківський національний університет радіоелектроніки,  
професор кафедри штучного інтелекту, м. Харків.

Офіційні опоненти: доктор технічних наук, професор  
Асєєв Георгій Георгійович,  
Харківська державна академія культури,  
завідувач кафедри інформаційних технологій;

доктор технічних наук, доцент  
Шостак Ігор Володимирович,  
Національний аерокосмічний університет імені  
М. Є. Жуковського, професор кафедри інженерії  
програмного забезпечення.

Захист відбудеться «19» червня 2013 р. о 15<sup>00</sup> годині на засіданні спеціалізованої вченої ради Д 64.052.01 у Харківському національному університеті радіоелектроніки за адресою: 61166, м. Харків, пр. Леніна, 14.

З дисертацією можна ознайомитися у бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Леніна, 14.

Автореферат розісланий «17» травня 2013 р.

Вчений секретар  
спеціалізованої вченої ради

Є.І. Литвинова

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** В основі онтологічної моделі представлення знань про об'єкти деякої предметної галузі лежить процес створення переліку логічних правил, що описують взаємодію заданих об'єктів через їх властивості. Онтологія дозволяє зберігати чисельні характеристики параметрів об'єктів реального світу. Але можливість опису взаємодії між такими параметрами відсутня. Таким чином існування будь-якого знання про відношення між числовими характеристиками об'єктів онтології, що формально представлене математичним відношенням, не можливо зафіксувати засобами Semantic Web.

Знання про математичні залежності між характеристиками об'єктів онтології предметної галузі може вирішити низку існуючих в онтологічному інжинірингу проблем: перевірка цілісності та несуперечності знань в онтології; отримання нових знань; класифікація об'єктів онтології згідно значень їх параметрів.

В дослідженнях опису математичного знання як одного з типів знань про відношення між об'єктами в онтологічних моделях найбільший внесок зробили вчені: Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosz, Mike Dean, Boris Motik, Bijan Parsia, Uli Sattler, Holger Knublauch.

На відміну від раніше пропонованих рішень щодо опису математичного знання між об'єктами онтологічної бази знань, в поданій роботі розглянуто розв'язання не тільки задачі розрахунку чисельних характеристик параметрів об'єктів онтологічної моделі, але й задачі створення формального опису – анування самого відношення в онтології та параметрів, які до нього входять.

**Зв'язок теми дисертації з планами наукових робіт.** Дисертаційна робота виконувалась згідно з науково-дослідним планом кафедри штучного інтелекту Харківського національного університету радіоелектроніки. Дослідження проводилися в межах держбюджетних тем: №195 «Розробка теоретичних засад, методів та моделей інтелектуальної обробки інформації та менеджменту знань у системах розподіленого штучного інтелекту» (№ ГР 0106U003286), №219 «Розробка Web-орієнтованої системи для підтримки процедур акредитації та ліцензування вищих навчальних закладів України» (№ ГР 0108U010139), №233 «Розробка системи підтримки семантичних запитів до онтологічної бази акредитації і ліцензування» (№ ГР 0109U001647). В межах вказаних тем здобувачем були розроблені та запропоновані модель семантичної анотації математичної залежності, опис семантичної анотації математичної залежності засобами мови RDFS, метод розширення семантики відношень в онтології предметної галузі за рахунок включення до неї знань про математичні залежності між об'єктами, модель інтеграції семантичної анотації математичної залежності в онтологічну базу знань, опис моделі інтеграції семантичної анотації математичної залежності в онтологію предметної галузі засобами мови OWL.

**Мета і задачі дослідження.** Метою дисертаційної роботи є розробка методів та моделей опису математичного знання, поданого у формі математичної залежності, удосконалення онтологічних баз знань шляхом розширення семантики відношень між об'єктами онтологій на основі знання про математичні залежності між ними.

У поданій дисертаційній роботі відповідно до мети дослідження вирішуються наступні задачі:

- виявлення основних принципів формального опису знань, поданих у вигляді математичного виразу, для побудови семантичної анотації математичної залежності;
- розробка моделі семантичної анотації математичної залежності, яка описує залежність як зовнішній щодо онтології предметної галузі інформаційний об'єкт;
- формалізація моделі семантичної анотації математичної залежності з використанням засобів мови RDFS;
- створення моделі інтеграції семантичної анотації математичної залежності до онтології предметної галузі та її формалізація засобами мов OWL та SWRL;
- розробка метода розширення семантики відношень в онтології предметної галузі шляхом додавання до опису відношень знань про математичні залежності між об'єктами онтології;
- експериментальне підтвердження можливості застосування розроблених метода та моделей для здійснення процедур логічного виведення з урахуванням доданих до онтології математичних залежностей, які описані семантичними анотаціями, на базі середовища розробки онтологічних моделей TopBraid Composer.

*Об'єктом дослідження* є створення формального опису знання про відношення, що представлені у вигляді математичних залежностей, між об'єктами предметної галузі в онтологічних базах знань.

*Предметом дослідження* є методи та моделі розширення семантики відношень між об'єктами онтології предметної галузі на основі інтеграції до неї семантичних анотацій математичних залежностей з використанням технологій онтологічного інжинірингу.

*Методами дослідження* є методи штучного інтелекту: теорія формальних мов та систем, що вирішує задачу формалізації семантичних характеристик математичної залежності та процедури її включення до онтологічної бази знань; технології онтологічного інжинірингу, застосовані для розробки моделі семантичної анотації математичної залежності та моделі розширення онтології предметної галузі новим знанням про відношення між об'єктами; технології Semantic Web, покладені в основу реалізації семантичної анотації математичної залежності та її інтеграції до онтології предметної галузі. Експериментальні дослідження проводилися в лабораторних умовах та на реальних об'єктах.

**Наукова новизна отриманих результатів.** В процесі вирішення поставлених задач в дисертаційній роботі були отримані такі наукові результати:

1. Вперше запропоновано формальну модель семантичної анотації математичної залежності як особливого типу знань, яка містить в собі семантичні характеристики математичного виразу та опис засобів його реалізації в інформаційному просторі, що дозволяє створювати семантичні описи зовнішніх по відношенню до онтології інформаційних об'єктів, які реалізують обчислення математичних виразів, та інтегрувати такі описи в онтології як математичні знання.

2. Вперше запропоновано формальну модель інтеграції нового знання, поданого у вигляді математичного виразу, в онтологію предметної галузі на основі застосування семантичної анотації математичної залежності, яка є сполучною

ланкою між цільовою онтологією предметної галузі та онтологією опису математичного знання. Це дозволяє розширити семантику відношень в онтології предметної галузі за рахунок використання формального знання про задане відношення, яке описане в термінах онтології математичного знання, що в свою чергу створює можливість використовувати дві інтегровані онтології як одне ціле, не впливаючи на їх життєвий цикл.

3. Удосконалено метод розширення семантики відношень між об'єктами онтологічної бази знань, що бере за основу метод інтеграції онтологічних структур – узгодження онтологій, який, на відміну від аналогів, передбачає уточнення семантичного опису відношень між об'єктами онтологічної структури у вигляді системи додаткових обмежень, що забезпечує можливість інтеграції онтологічних структур з різними чи непересічними контекстами застосування та дозволяє описувати об'єкти однієї онтології в термінах іншої.

**Практичне значення отриманих результатів.** Запропоновані моделі та метод можуть бути використані в побудові інтелектуальних інформаційних систем, призначених для накопичення об'єктів, що містять числові характеристики, відношення між якими встановлюються на основі математичних залежностей між значеннями таких характеристик.

Програмна реалізація інтерпретації розроблених моделей в удосконаленому середовищі онтологічного моделювання TopBraid Composer підвищує ефективність онтологічного інжинірингу, що може бути застосовано для опису специфічних предметних галузей, об'єкти яких характеризуються великою кількістю відношень між собою, дає наступні можливості:

- створення формального опису інформаційних об'єктів, що містять пов'язані числові характеристики;
- вирішення задачі перевірки цілісності та несуперечності онтологічних баз знань;
- створення механізмів перевірки достовірності інформації, даних та знань, що представлені в онтологічних базах знань;
- реалізація процедур автоматичного видобування знань в онтологічних системах на основі формального описаних правил взаємодії між об'єктами.

На основі розроблених в роботі формальної моделі семантичної анотації математичної залежності та методу розширення семантики відношень створено модуль перевірки достовірності вхідних даних для інформаційно-аналітичної системи зберігання статистичної інформації ТОВ “ІА “Інтерфакс-Україна” (акт впровадження від 24.04.2012 р.).

Розроблені моделі використано для опису показників діяльності ВНЗ в Web-орієнтованій системі для підтримки процедур акредитації та ліцензування вищих навчальних закладів України в межах держбюджетних тем: №195 «Розробка теоретичних засад, методів та моделей інтелектуальної обробки інформації та менеджменту знань у системах розподіленого штучного інтелекту» (№ ГР 0106U003286), №219 «Розробка Web-орієнтованої системи для підтримки процедур акредитації та ліцензування вищих навчальних закладів України» (№ ГР 0108U010139), №233 «Розробка системи підтримки семантичних запитів до

онтологічної бази акредитації і ліцензування» (№ ГР 0109U001647) (акт впровадження від 22.11.2012 р.).

Результати дисертаційної роботи також впроваджено в навчальному процесі кафедри штучного інтелекту Харківського національного університету радіоелектроніки у дисципліні «Інтелектуальні технології в Internet», «Системи розподіленого штучного інтелекту» (акт впровадження від 21.11.2012 р.).

**Особистий внесок здобувача.** Всі результати, що подані до захисту, отримані здобувачем особисто. В роботах, які опубліковано у співавторстві, здобувачу належать: [1] – процедури обчислення параметрів онтології для автоматизованого формування семантичної анотації документів семантичного репозиторію, інтеграція семантичної онтологічного порталу та семантичного репозиторію з використанням семантичної анотації документів; [2] – інтеграція онтологічної системи із зовнішніми Web-сервісами, що забезпечують можливість обробки даних водночас з декількох онтологічних моделей, концепція семантичного репозиторію для інтеграції онтологічної бази знань та зовнішніх інформаційних ресурсів; [5] – методологія функціонального розподілення ресурсів в онтологічній системі та розподілення інформаційної системи на рівні ресурсів з урахуванням характеру відношень між ними, метод розширення семантики відношень в онтології предметної галузі; [6] – модель формального опису ресурсів семантичного репозиторію, онтологія ресурсів документообігу в семантичному репозиторію; [7] – архітектура системи керування інформаційними ресурсами з їх семантичним описом у вигляді RDF-моделі, розробка інтерфейсу взаємодії системи з інтелектуальними агентами з використанням машини логічного виведення, що основана на логіці OWL; [10] – розробка концепції системи зберігання, класифікації та пошуку наукових статей з можливістю опису інформації, яка зберігається, в XML-форматі.

**Апробація результатів дисертації.** Основні положення та результати дисертаційної роботи були представлені, доповідались та обговорювались на міжнародних науково-технічних конференціях: на 8-му та 12-му міжнародному форумі «Радіоелектроніка та молодь у XXI сторіччі» (Харків, Україна, 2004 та 2008 рр.); на 11-ій Міжнародній науково-технічній конференції «Системний аналіз та інформаційні технології» (Київ, Україна, 2009 р.); на 6-ій Міжнародній науково-практичній конференції «Наука та соціальні проблеми суспільства: інформатизація та інформаційні технології» (Харків, Україна, 2011 р.).

**Публікації.** За результатами досліджень опубліковано 10 наукових праць: 6 статей (2 одноосібно) у наукових фахових виданнях України та 4 публікації (2 одноосібно) у збірках наукових праць, матеріалах, тезах доповідей міжнародних науково-технічних конференцій та форумів.

**Структура й обсяг дисертаційної роботи.** Дисертація складається із вступу, чотирьох розділів, висновків, переліку використаних джерел та додатків. Загальний обсяг дисертації становить 162 сторінки, з них 139 – обсяг основного тексту. Робота містить 32 рисунки, з яких 1 рисунок на окремій сторінці, 2 таблиці, 5 додатків на 12 сторінках та перелік використаних джерел із 106 найменувань на 11 сторінках.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і задачі дослідження, наведено відомості щодо наукової новизни отриманих у дисертації результатів, визначено їх практичну цінність, наведено відомості про апробацію та впровадження результатів.

**Перший розділ** містить огляд проблемної області і постановку задач дисертаційного дослідження. Детально розглянуто поняття семантичного анотування інформаційних ресурсів у середовищі World Wide Web, проведено розмежування понять анотування та семантичного анотування інформаційних ресурсів. Наведено визначення: інформаційний об'єкт – це представлення об'єкта предметної галузі в інформаційній системі, яке визначає його структуру, атрибути, обмеження цілісності та, можливо, поведінку; інформаційний простір – сукупність результатів діяльності людства, що представлені у вигляді самостійних інформаційних одиниць, семантичні характеристики яких дозволяють встановлювати зв'язок між ними.

Розглянуто принципи проектування баз знань та представлення інформації в них, можливість опису інформаційних ресурсів за допомогою онтологічного інжинірингу та побудови онтологій предметних галузей. Перевагою онтологій, як засобу представлення знань, є їх формальна структура, що спрощує їх комп'ютерну обробку.

Розглянуто існуючі технології формального опису математичних залежностей: Mizar Language, мови TeX, LaTeX, OpenMath, MathML, OMDoc, Semantic Web Rule Language (SWRL).

У результаті проведеного аналізу сформульовано основні завдання наукового дослідження, викладеного в дисертаційній роботі.

У **другому розділі** запропоновано модель семантичної анотації математичної залежності. Для виявлення характеристик математичної залежності вважається, що вона існує у деякому заданому середовищі. Середовище визначено як перелік об'єктів реального світу заданої предметної галузі, що взаємодіють між собою. Причому така взаємодія є функціональною з точки зору спостерігача та може бути описана математичною формулою. Математична залежність визначається двома пов'язаними параметрами середовища, в якому вона існує: 1) перелік умов, що описують стан середовища, в якому існує задана математична залежність; 2) механізм, що пов'язує формальний опис об'єктів середовища, і, таким чином, забезпечує реалізацію заданої залежності.

Для встановлення семантичних характеристик математичної залежності необхідно розглянути семантичні характеристики, що визначають середовище заданої залежності та механізм її реалізації.

Створення семантичної анотації математичної залежності передбачає необхідність опису об'єктів середовища, які входять до відношення, що описане заданою залежністю. Також повинен бути реалізований опис всіх параметрів об'єктів, що є в формулі, яка описує задану залежність.

З визначення семантичної анотації маємо, що її структура та формат включених до неї об'єктів залежать від призначення анотації та структури інформаційного об'єкта, який вона описує.

Для формування моделі семантичної анотації математичної залежності введемо систему позначень:

- 1)  $\Omega$  – знання про поняття та залежності в предметній галузі, характеристики яких можуть бути виражені математично;
- 2)  $O$  – онтологія, яка містить терміни та відношення, необхідні для семантичного опису математичних залежностей;
- 3)  $\varphi$  – математична залежність, яка визначає відношення між об'єктами  $\Omega$ , що представлені в онтології  $O$ ;
- 4)  $\psi$  – запис математичного виразу, що представляє залежність  $\varphi$ ;
- 5)  $\sigma$  – зовнішній сервіс, який реалізує обчислення математичного виразу  $\psi$ ;
- 6)  $f$  – анотація математичної залежності  $\varphi$ , яка реалізована зовнішнім сервісом  $\sigma$  та виражена термінами онтології  $O$ .

Для представлення математичної залежності у вигляді семантичної анотації інформаційного об'єкту з урахуванням зазначених вимог була розроблена модель (рис. 1), що відображає знання про математичну залежність та сервіс, що її реалізовує, в семантичну анотацію.

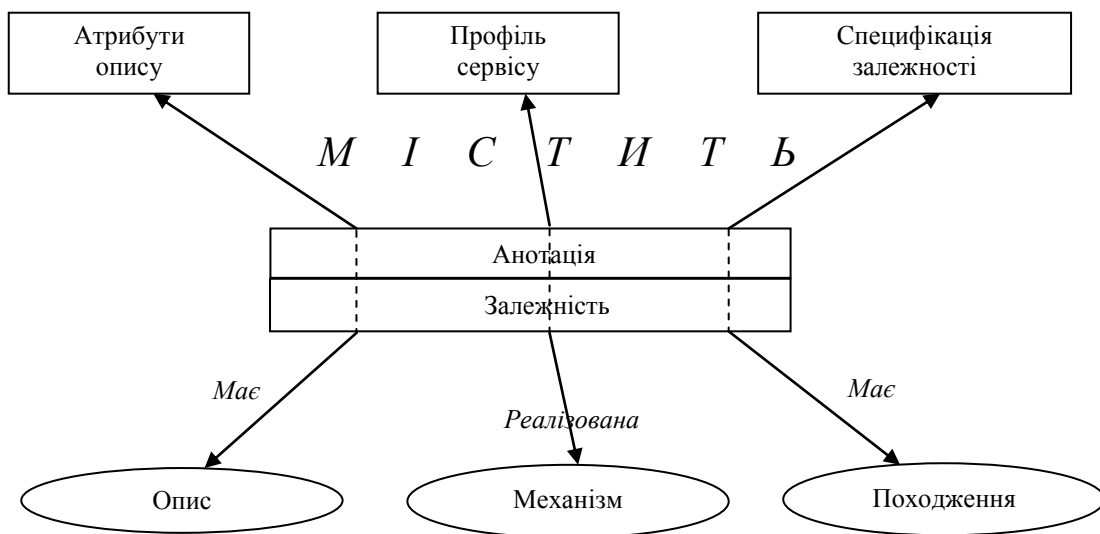


Рис. 1. Модель відображення математичної залежності в її семантичну анотацію

Для визначення структури інформаційного об'єкту  $f$  визначено таку систему позначень, що відображає семантику математичної залежності  $\varphi$ :

- 1)  $P$  – множина властивостей заданої онтології  $O$ , що відповідають всім параметрам виразу залежності  $\varphi$ ;
- 2)  $C$  – перелік класів заданої онтології  $O$ , що є поєднанням та включає:
  - $C_L$  – множина класів, деякі відносини яких визначаються математичною залежністю  $\varphi$ ;
  - $C_D$  – множина класів, що відповідають області визначення застосування параметрів виразу  $\varphi$ ;



- $C_R$  – множина класів, що визначають допустимі значення параметрів виразу  $\psi$ ;

3)  $X$  – множина нерівностей, що задають діапазони значень параметрів. Такі діапазони можуть бути виражені явно, в числовій формі, або за допомогою інших параметрів виразу  $\psi$ .

Множина нерівностей  $X$  специфікує вимоги, при яких математичний вираз  $\psi$  має сенс в заданій предметній галузі.

Семантичне визначення параметрів  $P$  та  $C$  під час анотування математичної залежності здійснюється за допомогою визначення їх відповідності знанню  $\Omega$ , яке формально описане в онтології предметної галузі  $O$ .

Важливим елементом опису математичної залежності  $\varphi$ , що необхідно включити до семантичної анотації  $f$ , є анотація  $A$  – анотація математичної залежності, яка представлена множиною впорядкованих пар типу  $\langle n, v \rangle$ , де

- $n$  – терм анотації;
- $v$  – значення терму анотації.

Фактично, всі розглянуті параметри анотації математичної залежності можуть бути представлені кортежем множин

$$f = \{ P, C, X, A \} \quad (1)$$

Таким чином семантична анотація математичної залежності містить в собі опис того, до чого саме може бути застосована математична залежність в межах заданої предметної галузі. Це надає можливість розглядати таку математичну залежність як додаткове знання, яке віднесене до конкретного об'єкту заданої онтологічної структури.

Під профілем математичної залежності розуміється опис схеми зовнішнього сервісу  $\sigma$  та елементів анотації, необхідних для зв'язку онтологічної моделі з термінами сервісу. В цьому разі схема сервісу містить в собі символічний опис математичної залежності, представлений у вигляді математичного виразу, а також атрибутів, що необхідні для ідентифікації потрібного сервісу.

Для опису профілю математичної залежності задано таку систему визначень: нехай,  $s$  – профіль математичної залежності  $\varphi$ , тоді

$$s = \{ M, g, t, d \}, \text{ де} \quad (2)$$

$M$  – перелік співвідношень властивостей  $p_i$  з множини  $P$  до параметрів  $q_i$  математичного виразу  $\psi$ , який має вигляд впорядкованої пари  $\langle p, q \rangle$ , де

- $p$  – ідентифікатор властивості;
- $q$  – символічний запис параметру;
- $g$  – формальний запис математичного виразу;
- $t$  – формат сервісу  $\sigma$ , заданий деяким протоколом доступу до нього;
- $d$  – адреса або ідентифікатор сервісу  $\sigma$ .

Модель семантичної анотації математичної залежності повинна бути реалізована в термінах RDFS. Таким чином для формального запису анотації необхідно розробити моделі деяких класів (*rdfs:Class*) та властивостей (*rdf:Property*),

які будуть відповідати частинам анотації, що описується. Беручи до уваги, що формальний опис семантичної анотації представлено за допомогою ресурсів, якими вона оперує, є виправданим формування моделі анотації, що представлена класами:

- 1) математична залежність *Expression*;
- 2) математичний вираз *Blackbox* та відповідний до нього зовнішній сервіс;
- 3) змінні *Variable*, які використовуює математична залежність.

Властивості класу *Expression* визначено наступним чином:

1) *link* – властивість, що вказує на клас, між параметрами якого встановлюється задана математична залежність, чи перелік класів, параметри яких включені в аналітичний вираз заданої математичної залежності; семантично така властивість відповідає параметру  $C_L$  в моделі семантичної анотації  $f$ ;

2) *operate* – множинна властивість, що визначає перелік змінних, представлених в класі *Variable*. При цьому екземпляри класу *Variable* мають наступні характеристики:

– властивості онтології (*property*), які є в заданій математичній залежності, семантично відповідають множині  $P$  в моделі анотації  $f$  та переліку всіх  $p$  в множині  $M$  моделі профілю  $s$ ;

– символічний запис (*symbol*) змінних, що входять до аналітичного опису математичної залежності, яка представлений у вигляді елементарної SWRL-конструкції; семантично відповідає переліку всіх  $q$  у множині  $M$  моделі профілю  $s$ ;

– область визначення (*domain*) заданих властивостей в межах опису заданого математичного виразу, що семантично відповідає множині  $C_D$  в моделі семантичної анотації  $f$ ;

– область допустимих значень (*range*) заданих властивостей, що можуть бути оброблені зовнішнім сервісом, що реалізує інтерпретацію заданої математичної залежності; семантично відповідає множині  $C_R$  в моделі семантичної анотації  $f$ .

Таким чином створення класу *Variable* дозволить уточнити властивості та область дії елементів заданої онтології, до яких застосовується обрана математична залежність, що по суті є новим знанням в онтології предметної галузі;

3) *scope* – властивість задається значенням *swrl:AtomList* та представляє атоми, що містять математичні нерівності, які пов'язують відповідні до математичної залежності змінні. Основним переліком атомів є процедури *Built-In*, які описують арифметичні перетворення та нерівності. Але за необхідністю ця множина може бути розширена згідно існуючої специфікації мови SWRL. Така послідовність атомів семантично відповідає множині  $X$  в моделі семантичної анотації  $f$ ;

4) *rdfs:isDefinedBy* – властивість схеми мови RDFS, що визначає екземпляр класу *Blackbox*. Цей клас вводиться для представлення профілю  $s$  заданої математичної залежності, яка розглядається з точки зору заданої онтології як «чорна скриня»;

5) довільні анотаційні властивості словників RDFS та Dublin Core, які відповідають сформованій анотації  $A$  з моделі семантичної анотації  $f$ , де  $n$  – предикат словника, а  $v$  – його літеральне значення.

Слід зазначити, що для формування анотації  $A$  математичної залежності, яка містить характеристики цієї залежності як зовнішнього інформаційного об'єкту,

дозволяється використовувати будь-які властивості Dublin Core, окрім *format* і *source*, які мають спеціальну інтерпретацію.

Визначимо клас *Blackbox* такими властивостями:

1) властивості схеми Dublin Core – *format* та *source*, задані у вигляді, який може бути коректно інтерпретований сервісом, що реалізує математичну залежність; семантично відповідає параметрам  $t$  та  $d$  профілю  $s$  моделі семантичної анотації  $f$ ;

2) *formula* – властивість, що містить візуальне представлення математичної залежності, створене за допомогою мови MathML; семантично відповідає параметру  $g$  профілю  $s$ ;

3) *argument* – перелік символів, що представляють змінні математичного виразу, тобто аргументи зовнішнього сервісу, які записані у вигляді переліку конструкцій *swrl:Variable* мови SWRL; семантично відповідають переліку всіх  $q$  множини  $M$  моделі профілю  $s$ .

Розроблювана онтологія містить опис знання про терми, які використовуються у визначенні математичного виразу, а відповідна йому семантична анотація вказує лише на засіб реалізації математичної залежності, що описана в онтології.

Виходячи з замкнутості контуру відношень між математичною залежністю, відповідною онтологією понять та семантичною анотацією заданої залежності, що представлено на рис. 2, слід зазначити, що процеси створення онтології та анотації взаємопов'язані та ітеративні.

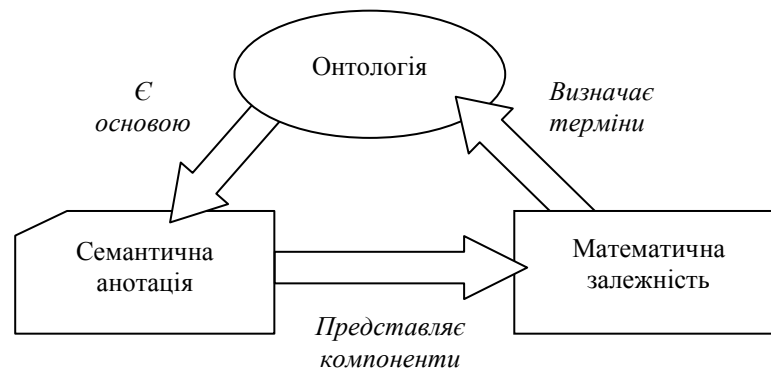


Рис. 2. Відношення між математичною залежністю та її семантичною анотацією

Визначений підхід до створення семантичної анотації математичної залежності відповідає загальним принципам онтологічного інжинірингу та фактично є процесом вивчення та опису об'єктів та їх взаємодії в контексті онтології.

У **третьому розділі** запропоновано метод розширення семантики відношень в онтології предметної галузі та модель інтеграції семантичної анотації математичної залежності до онтології предметної галузі. При цьому будь-яке розширення онтології предметної галузі слід розглядати як уточнення та доповнення знань, які в ній представлені.

У розглянутому підході узгодження онтологій, який було обрано за основу інтеграції онтологій, інтеграція відбувається шляхом створення посилань між пов'язуваними об'єктами двох інтегрованих онтологій. Але такий підхід передбачає статичність інтегрованих онтологічних моделей.

Запропонований в дисертаційній роботі метод розширення семантики відношень в онтології предметної галузі пропонує створення деякої системи правил, які будуть пов'язувати дві онтологічні структури, що інтегруються. Це надає можливість не змінювати вже існуючі онтології, а також не створювати додаткових онтологічних моделей, які повинні описувати результат інтеграції похідних онтологій. Це, в свою чергу, надає можливість здійснювати ефективну взаємодію похідних онтологій під час усього життєвого циклу без додаткової роботи з організації інтеграції.

Для докладного опису метода розширення семантики відношень в онтології предметної галузі за допомогою інтеграції анотації математичних залежностей уточнимо систему позначень, що була запропонована в Розділі 2, наступним:

- 1)  $\Theta$  – знання про деяку предметну галузь, для об'єктів якої характерні математичні залежності знання  $\Omega$ ;
- 2)  $Q$  – онтологія предметної галузі, яка виражає знання  $\Theta$ ;
- 3)  $Q'$  – розширена онтологія  $Q$ , яка містить знання  $\Omega$  про природу математичних залежностей;
- 4)  $\Phi$  – модель інтеграції онтології  $O$  в онтологію  $Q$  за допомогою семантичної анотації  $f$  математичної залежності  $\varphi$ .

Згідно визначеним об'єктам, які приймають участь в процесі розширення семантики відношень в онтології за допомогою опису математичних залежностей, інтеграцію нового знання до цільової онтології можна представити функціональною діаграмою, яка представлена на рис. 3.

Інтеграція семантичної анотації зовнішнього інформаційного об'єкту та заданої онтології не є поєднанням двох онтологічних моделей – це розширення існуючої моделі новим знанням (рис. 4).

Розглянута інтеграція зовнішнього інформаційного об'єкту, у ролі якого є семантична анотація математичної залежності  $f$ , до цільової онтології  $Q'$  повинна проводитися шляхом імпорту такого об'єкту до онтології предметної галузі з наступним створенням правил, що забезпечать взаємодію похідних онтологій  $Q$  та  $O$ .



Рис. 3. Функціональна діаграма, що описує процес включення математичного знання до онтології предметної галузі

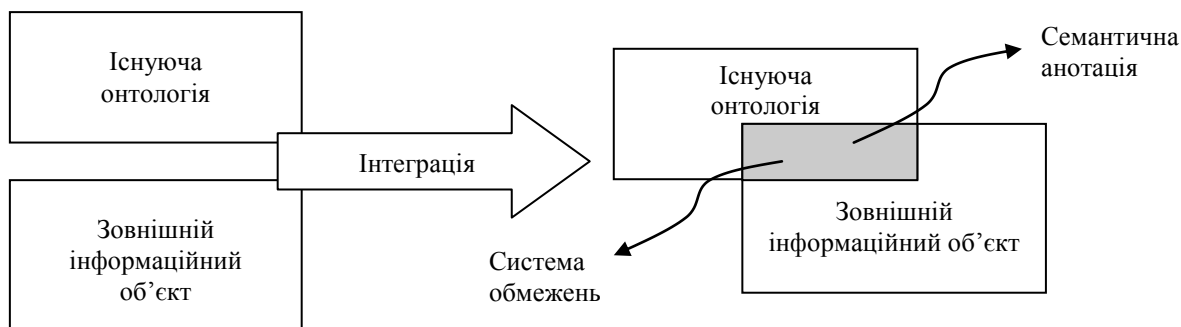


Рис. 4. Система інтеграції семантичної анотації предметної галузі та семантичної анотації зовнішнього інформаційного об'єкту

Таким чином, інтеграція онтологій відбувається на рівні понять, що дозволяє виключити реалізацію класичних механізмів інтеграції. Але при цьому можна стверджувати, що відбувається включення до похідної онтології  $Q$  деяких, необхідних для вирішення конкретних задач, знань з онтології  $O$ .

Модель інтеграції  $\Phi$  має вигляд:

$$\Phi = \{F, R, E\}, \text{ де} \quad (3)$$

$F$  – множина семантичних анотацій  $f$ , які представляють математичні залежності  $\varphi$  та виражають математичні відношення, що характерні для заданого зв'язку чи екземпляру онтології  $Q$ ;

$R$  – набір продукційних правил, що визначають зв'язок між екземплярами онтології  $Q$  та відповідними поняттями онтології  $O$ , виду  $H \rightarrow B$ , де

- $H$  – логічна формула, яка визначає екземпляри онтології  $Q$ , пов'язані за допомогою математичної залежності  $\varphi$ ,
- $B$  – логічна формул, що співвідносить властивості екземплярів онтології  $Q$ , що були обрані в  $H$ , з параметрами математичної залежності  $\varphi$ , представленими в анотації  $f$ ;

$E$  – специфікація обмежень, за допомогою яких клас чи відношення в онтології  $Q$  виражаються переліком залежностей, які представлені анотаціями в  $F$ .

Інтерпретація запропонованої моделі  $\Phi$  полягає у наступному:

1) при знаходженні в множині  $R$  такого правила, що зіставляє всі змінні деякої семантичної анотації математичної залежності  $f_i \in F$  з властивостями класів заданої онтології  $Q$ , можна вважати, що відповідна залежність  $\varphi_i$  визначає специфікацію обмежень  $E$ ;

2) якщо всі значення параметрів відповідних екземплярів визначені через правила  $R$  в заданій онтології  $Q$ , та математичний вираз  $\psi_i$  анотації  $f_i \in F$  при цьому є істинним, то всі специфікації обмежень  $E$  є істинними;

3) якщо один або декілька параметрів екземплярів класів, обраних згідно правилам  $R$  для відповідної анотації  $f_i \in F$ , не визначені в онтології  $Q$ , то при виконанні множини обмежень  $E$  відсутні значення можуть бути отримані шляхом обчислення функції  $\psi_i$  згідно анотації математичної залежності  $f_i$ .

Для реалізації семантичної моделі інтеграції  $\Phi$  математичної залежності в онтологічну базу знань пропонується розширити моделі семантичного опису OWL та SWRL метамоделлю на основі RDFS, яка містить наступні компоненти:

1) клас *Dependence*, який реалізує модель інтеграції  $\Phi$  та визначає перелік математичних залежностей та логічних зв'язків, який є підкласом класу *swrl:Imp*. Згідно моделі SWRL в ньому визначені предикати *swrl:body* та *swrl:head*, які відповідають правилам  $R$  в моделі інтеграції  $\Phi$  та являють собою множини  $H$  та  $B$ . Перелік атомів в *swrl:body* представляє логічну формулу, що визначає окремі екземпляри заданої онтології  $Q$ , властивості яких можуть бути пов'язані за допомогою семантичних анотацій математичних залежностей  $F$ . Атоми, що записані в *swrl:head*, визначають співвідношення властивостей екземплярів онтології  $Q$ , що були представлені в  $H$ , з параметрами семантичними анотаціями математичних залежностей, представлених в  $F$ ;

2) властивість *express*, визначена для класу *Dependence*, яка представляє множину семантичних анотацій математичних залежностей  $F$ , що входять у відповідну моль  $\Phi$ ;

3) властивість *implicate*, діапазон значень якої визначений класом *Dependence*, яка описує частину системи обмежень  $E$  (друга частина обмежень реалізована з використанням *owl:Restriction* моделі OWL) в застосуванні до існуючих класів *owl:Class* та властивостей *owl:ObjectProperty* заданої онтології  $Q$ . Властивість призначено для фіксації в межах моделі  $\Phi$  логічних або математичних відношень, що мають на увазі, у відповідних елементах похідної онтології  $Q$ .

Результатом реалізації моделі інтеграції є деякий сполучний ресурс та його характеристики, що поєднують між собою поняття та властивості двох похідних онтологій. При цьому модифікація цільової онтології шляхом додавання до неї опису обмежень, згідно запропонованому методу інтеграції, має на увазі зв'язування отриманого ресурсу, тобто екземпляру класу *Dependence*, з онтологією предметної галузі.

У **четвертому розділі** наведено результати моделювання запропонованих у дисертаційній роботі моделей і метода семантичного опису математичних залежностей та розв'язання практичних задач за їхньою допомогою.

Експериментальна реалізація розробленого методу інтеграції полягає в підключенні ядра обчислювального математичного пакета Mathematica, який представлено відповідним Java-адаптером, до системи TopBraid Composer для обробки та обчислення математичних виразів з використанням модуля розширення Eclipse – ExtMath (рис. 5). Зв'язок з пакетом Mathematica реалізований двома SPARQL-функціями, що використані в SPIN-правилах.

При підключенні розробленого модуля ExtMath до середовища TopBraid Composer вбудований інструмент логічного виведення TopSPIN доповнюється наступними можливостями обробки SPIN-правил:

1) можливість розв'язувати рівняння та системи рівнянь, які містять обмеження у вигляді нерівностей, для знаходження однієї змінної (за допомогою функції *Solve*);

2) можливість перевіряти істинність рівняння при наявності всіх відомих величин (за допомогою функції *Test*).

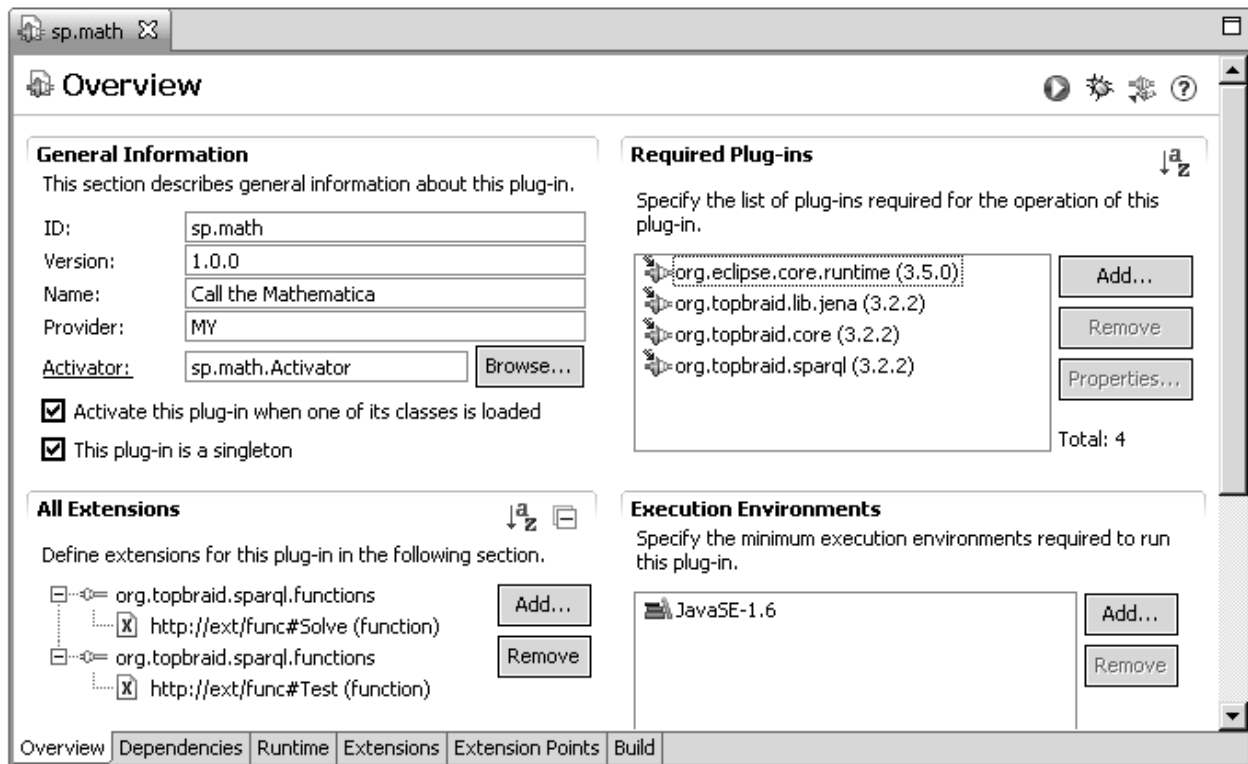


Рис. 5. Проект модуля розширення середовища Eclipse, що реалізує дві додаткові SPARQL-функції

Запропоновані в роботі модель семантичної анотації та модель інтеграції не є SPIN-правилами. Це визначає необхідність створення додаткового механізму перетворення розроблених семантичних моделей (модель семантичної анотації  $f$  та модель інтеграції математичної залежності  $\Phi$ ) в RDF-графи відповідних SPIN-правил – *DepsImpl*. Для реалізації такого перетворення була використана мова Perl, за допомогою якої була розроблена скриптова програма – *ConvMath*, що виконує синтаксичне перетворення актуальних елементів моделі інтеграції, які вилучені з онтології OWL за допомогою виконання SPARQL-запиту – *DepsQuery*.

Результатом експериментальних досліджень стало розширення можливостей вбудованої до середовища розробки онтології TopBraid Composer машини логічного виведення TopSPIN. Запропонований підхід забезпечує можливість виконувати обчислення, що основані на розроблених моделях та методи, в онтологічній моделі предметної галузі.

У **висновках** сформульовано наукові та практичні результати, що їх одержано в дисертаційній роботі.

**Додаток** містить копії документів про впровадження та практичне застосування результатів, отриманих у дисертаційній роботі.

## ВИСНОВКИ

В дисертаційній роботі наведено результати, які відповідно до мети дослідження є вирішенням науково-технічної задачі удосконалення можливостей

онтологічних баз знань шляхом інтеграції моделей, що розширюють семантику відношень між об'єктами онтології за рахунок включення знання про математичні залежності між ними. Проведені дослідження дозволили отримати такі результати:

1. Вперше запропоновано формальну модель семантичної анотації математичної залежності. Для створення семантичної анотації математичної залежності було виявлено семантичні характеристики математичного виразу, що описує задану залежність. Виявлені семантичні характеристики математичної залежності описуються за допомогою стандартів Semantic Web в термінах онтології предметної галузі, до якої буде інтегрована задана семантична анотація.

2. Вперше запропоновано модель інтеграції нового знання, що представлене у вигляді математичного виразу, до онтології предметної галузі на основі використання семантичної анотації математично залежності, яка виступає в ролі зв'язку між цільовою онтологією предметної галузі та онтологією опису математичного знання. Це створює можливість використовувати дві інтегровані онтології як одне ціле, не впливаючи на їх життєвий цикл.

3. Удосконалено метод розширення семантики відношень між об'єктами онтології предметної галузі, що базується на методі інтеграції онтологічних структур – узгодженні онтологій. На відміну від аналогів запропонований метод передбачає уточнення семантичного опису відношень між об'єктами онтологічної структури у вигляді системи додаткових обмежень. Це дозволяє описувати об'єкти однієї онтології в термінах іншої, а також забезпечує можливість інтеграції онтологічних структур з різними чи не пересічними контекстами застосування.

4. Запропоновані метод та моделі формального опису математичного знання знайшли свою практичну реалізацію при розширенні мов OWL та RDFS в процесі удосконалення можливостей середовища онтологічного моделювання TopBraid Composer завдяки зв'язку з математичним пакетом Mathematica за рахунок додавання можливості оброблювати запропоновані моделі на експериментальних онтологіях предметних галузей. Це підтвердило можливість реалізації процедур автоматичного видобування знань в онтологічних системах на основі формально описаних правил взаємозв'язку між параметрами об'єктів.

В роботі розроблено підхід до інтеграції онтологічних моделей, що не потребує модифікації похідних онтологій та забезпечує їх подальшу взаємодію на протязі усього життєвого циклу. В основі такого підходу лежить семантична анотація математичної залежності, яка є сполучною ланкою між онтологією предметної галузі та онтологією, що містить опис математичних об'єктів та залежностей. В цьому випадку кожна семантична анотація розширює знання про зв'язок між об'єктами онтології предметної галузі за рахунок знання, що представлене в онтології математичних об'єктів та виразів.

Результати дисертаційної роботи були використані під час виконання НДР, впроваджені у навчальному процесі кафедри штучного інтелекту Харківського національного університету радіоелектроніки у дисципліни «Інтелектуальні технології в Internet», «Системи розподіленого штучного інтелекту», а також використані для розробки і реалізації інформаційно-аналітичної системи зберігання статистичної інформації ТОВ «ІА "Інтерфакс-Україна"».



## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Рябова Н.В. Прототип семантического репозитория документов распределенного онтологического портала / Н.В. Рябова, А.Ю. Шевченко, М.В. Головянко, Н.А. Волошина, А.А. Воскобойникова, М.В. Белоиваненко // Восточно-Европейский журнал передовых технологий. – Харьков, 2012. – Вып. 4/2(58). С. 56-62
2. Рябова Н.В. Разработка обобщенной аппаратно-программной архитектуры распределенной версии онтологического портала / Н.В. Рябова, А.Ю. Шевченко, М.В. Головянко, Н.А. Волошина, О.В. Шубкина, А.А. Воскобойникова, М.В. Белоиваненко // Восточно-Европейский журнал передовых технологий. – Харьков, 2012. – Вып. 3/11(57). С. 42-50
3. Белоиваненко М.В. Модель интеграции семантической аннотации математической зависимости в онтологические базы знаний / М.В. Белоиваненко // Восточно-Европейский журнал передовых технологий. – Харьков, 2012. – Вып. 2/2(56). С. 54-57
4. Белоиваненко М.В. Семантическая аннотация математической функции в системах интеллектуальной обработки информации / М.В. Белоиваненко // Бионика интеллекта. – Харьков, 2011. – Вып. 1(75). С. 66-69
5. Белоиваненко М.В. Розробка концепції розподілення інформаційної системи, яка побудована на онтологічній базі знань / М.В. Белоиваненко, Г.А. Воскобойникова // Системи обробки інформації. – Харків, 2010. – Вип. 1(82). С. 2–7.
6. Рябова Н.В. Розробка прототипу семантичного репозиторію документів з можливістю контролю версій та його інтеграція з розподіленим онтологічним порталом менеджменту освітніх та наукових ресурсів МОН України / Н.В. Рябова, А.Ю. Шевченко, Г.А. Воскобойникова, М.В. Белоиваненко, М.В. Головянко, Н.А. Волошина, О.В. Шубкина // Інформаційні технології в освіті. – Херсон, 2010. – Вип. 7. С. 117-126.
7. Белоиваненко М.В. Онтологический подход в системах управления Web-контентом / М.В. Белоиваненко, М.А. Шаламов // Материалы XI Международной научно-технической конференции «САИТ-2009». – К.: УНК «ИПСА» НТУУ «КПИ», 2009. – С. 265.
8. Белоиваненко М.В. Реализация алгебраических выражений в онтологическом портале МОНУ / М.В. Белоиваненко // Материалы 12-го Международного молодежного форума «Радиоэлектроника и молодежь в XXI веке»: – Харьков: ХНУРЕ, 2008. – С. 190.
9. Белоиваненко М.В. Интеграция семантической аннотации математической функции в онтологию предметной области / М.В. Белоиваненко // Материалы VI-й Международной научно-практической конференции «Наука и социальные проблемы общества: информатизация и информационные технологии». – Харьков, 2011. – С. 98-99.
10. Бачевский А.А. Перспективы развития систем накопления и обработки информации / А.А. Бачевский, М.В. Белоиваненко // Материалы 8-го Международного молодежного форума «Радиоэлектроника и молодежь в XXI веке»: – Харьков: ХНУРЕ, 2004. – С. 64.

## АНОТАЦІЯ

Білоіваненко М.В. Метод та моделі розширення семантики відношень онтологій за допомогою опису математичних залежностей між параметрами об'єктів. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук зі спеціальності 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки МОН України, Харків, 2013.

Дисертація присвячена дослідженню та розробці методів розширення функціональних можливостей онтологічних баз знань шляхом додавання до них нових, формально представлених у вигляді математичних залежностей, знань про взаємозв'язок між об'єктами.

Метою дисертаційної роботи є розробка методів та моделей опису математичного знання, удосконалення онтологічних баз знань шляхом розширення семантики відношень між об'єктами онтології на основі знання про математичні залежності між ними. Об'єктом дослідження є створення формального опису знання про відношення, що представлені у вигляді математичних залежностей, між об'єктами предметної галузі в онтологічних базах знань. Предметом дослідження є методи та моделі розширення семантики відношень між об'єктами онтології предметної галузі на основі інтеграції до неї семантичних анотацій математичних залежностей з використанням технологій онтологічного інжинірингу.

Вперше запропоновано формальну модель семантичної анотації математичної залежності та метод розширення семантики відношень між об'єктами онтології предметної галузі, який є удосконаленням метода інтеграції онтологій – узгодження онтологій. Вперше запропоновано формальну модель інтеграції нового знання, що представлено у вигляді математичного виразу, до онтології предметної галузі на основі використання семантичної анотації математичної залежності, яка виступає в ролі зв'язку між цільовою онтологією предметної галузі та онтологією опису математичного знання. Запропоновані метод та моделі формального опису математичного знання знайшли свою практичну реалізацію при розширенні можливостей мов Semantic Web – OWL та RDFS, а також при удосконаленні можливостей середовища онтологічного моделювання TopBraid Composer за рахунок додання до нього можливості оброблювати запропоновані моделі за рахунок обчислювальних можливостей пакету Mathematica.

**Ключові слова:** інформаційний ресурс, онтологічна база знань, математичне відношення, семантична анотація, технології Semantic Web, інтеграція онтологій.

## АННОТАЦИЯ

Белоиваненко М.В. Метод и модели расширения семантики отношений в онтологии с помощью описания математических зависимостей между параметрами объектов. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, МОН Украины,

Харьков, 2013.

Диссертация посвящена исследованию и разработке методов расширения функциональных возможностей онтологических баз знаний путем добавления в них новых, формально представленных в виде математических зависимостей, знаний о взаимосвязях между объектами.

Целью диссертационной работы является разработка методов и моделей описания математического знания, представленного в форме математической зависимости, усовершенствование онтологических баз знаний путем расширения семантики связей между объектами онтологии на основе знания о математических зависимостях между ними.

Объектом исследования является создание формального описания знания об отношениях, которые представлены в виде математических зависимостей, между объектами предметной области в онтологических базах знаний.

Предметом исследования являются методы и модели расширения семантики отношений между объектами онтологии предметной области на основе интеграции в нее семантических аннотаций математических зависимостей с использованием технологий онтологического инжиниринга.

В диссертационной работе подробно рассмотрено применение онтологических баз знаний как инструмента представления и обработки информационных объектов, описывающих объекты реального мира. Приведен обзор средств аннотирования математической информации. Рассмотрена задача аннотирования математических выражений как особого типа знаний, которые раскрывают характер отношений между объектами реального мира.

Во втором разделе рассмотрена необходимость создания семантической аннотации математической зависимости как внешнего по отношению к онтологии предметной области информационного объекта, которая составлена в терминах онтологии, описывающей порядок формирования заданных математических выражений.

Впервые предложена формальная модель семантической аннотации математической зависимости, которая содержит в себе семантические характеристики математического выражения и средств его реализации в информационном пространстве, а также описывает знания об отношениях между объектами предметной области. Данная модель описывает такие аспекты математической зависимости, которые явно характеризуют механизм, связывающий объекты онтологии. Т.о. знание о предметной области дополняется новым знанием о природе связей между входящими в нее объектами.

В третьем разделе впервые предложена модель дополнения онтологии предметной области новыми знаниями путем интеграции в нее внешнего информационного объекта в виде семантической аннотации математической зависимости, которая расширяет характеристики связей между объектами онтологии.

Детально рассмотрена семантическая интеграция внешнего информационного объекта, которая предполагает создание ссылки от заданной онтологии к семантической аннотации посредством создания правил и ограничений, что предполагает модификацию онтологии путем добавления специальных выражений в

описание сущностей. Данный подход позволяет включить новый информационный объект, вне зависимости от его контекста, в онтологию предметной области, не нарушив целостность её обработки.

Для реализации модели интеграции аннотации математической зависимости в онтологическую базу знаний были расширены языки OWL и SWRL разработанной метамоделью семантической аннотации на основе модели RDFS.

В четвертом разделе приведены результаты моделирования предложенных моделей и метода семантического описания математических зависимостей и решения практических задач с их помощью. Экспериментальная реализация разработанного метода интеграции заключается в подключении ядра вычислительного пакета Mathematica к системе TopBraid Composer для обработки математических зависимостей между объектами предметной области встроенной в среду разработки онтологий машиной логического вывода TopSPIN. Предложенный подход обеспечивает возможность выполнять вычисления, описания которых включены в онтологическую модель предметной области.

**Ключевые слова:** информационный ресурс, онтологическая база знаний, математическое отношение, семантическая аннотация, технологии Semantic Web, интеграция онтологий.

## ABSTRACT

Biloivanenko M.V. Method and models expanding semantics of the ontology relationships by describing the mathematical dependencies between the parameters of objects. – Manuscript.

Thesis for the candidate's degree in technical sciences, specialty 05.13.23 – systems and tools of artificial intelligence. – Kharkiv National University of Radioelectronics, Kharkiv, 2013.

The thesis is devoted to development the method and models for improvements ontological knowledge base capabilities. Through the development of model for the semantic annotation of mathematical dependencies obtained the ability to create formally description a new knowledge type – mathematical knowledge. The development expansion of the semantic relations in ontology method based-on semantic annotation is clarifies the constrain system on the values ontology objects characteristics.

The proposed formal description of the mathematical knowledge found its practical implementation in the expansion declarative opportunities of OWL and RDFS Semantic Web languages. Experimental implementation of proposed integration method is the connection kernel of Mathematica computing system with environment of ontological modeling TopBraid Composer for processing mathematical relationships between objects with integrated ontology inference engine TopSPIN.

This approach to the integration of the ontological models does not require modification of original ontologies and provides them further interaction throughout their life cycle.

**Key words:** information resource, ontological knowledge base, mathematical dependence, semantic annotation, Semantic Web technology, ontology integration.

Відповідальний випусковий О.Г. Руденко

Підп. до друку 16.05.13. Формат  $60 \times 84^{1/16}$ . Спосіб друку – ризографія.  
Умов. друк. арк. 1,2. Облік. вид. арк. 1,0. Тираж 100 прим. Зам. № 2-409.

Україна, 61166 Харків, просп. Леніна, 14, ХНУРЕ

Віддруковано в навчально-науковому  
видавничо-поліграфічному центрі ХНУРЕ.  
Україна, 61166 Харків, просп. Леніна, 14