DOI: https://doi.org/10.15407/kvt192.02.023

УДК 681.3.

В.І. ГРИЦЕНКО[1], член-кореспондент НАН України,
директор Міжнародного науково-навчального центру
інформаційних технологій та систем НАН України та МОН України
e-mail: vig@irtc.org.ua
А.Я. ГЛАДУН[1], канд.техн.наук, старш. наук. співроб.,
відд. комплексних досліджень інформаційних технологій та систем
email: glanat@yahoo.com
Ю.В. РОГУШИНА[2], канд.фіз.-мат. наук, старш. наук. співроб.,
відд. автоматизованих інформаційних систем
ladanandraka2010@gmail.com
[1] Міжнародний науково-навчальний центр інформаційних
технологій та систем НАН України та МОН України,
пр. Академіка Глушкова, 40, м. Київ, 03187, Україна
[2] Інститут програмних систем НАН України,
пр. Академіка Глушкова, 40, м. Київ, 03187, Україна

МОДЕЛІ ТА МЕТОДИ ВИКОРИСТАННЯ СЕМАНТИЧНИХ WIKI-РЕСУРСІВ ЯК ДЖЕРЕЛА ЗНАНЬ ДЛЯ ПОПОВНЕННЯ ФОРМАЛЬНИХ ОНТОЛОГІЙ ПРЕДМЕТНИХ ОБЛАСТЕЙ

Розглянуто доцільність застосування онтологій для подання знань у системах штучного інтелекту, орієнтованих на функціонування у відкритому середовищі Web. Проаналізовано дослідження, пов'язані з побудовою формальних онтологій предметних областей, та засоби їх формалізації. Запропоновано формальну модель онтології, що уточнює наявні підходи та більш детально описує властивості та характеристики відношень між основними елементами онтології. Запропоновано використання такої моделі як основи для співставлення онтології з дескриптивними логіками. Подано метод поповнення онтології предметної області, який базується на запропонованій моделі та використанні семантично розмічених Wiki-ресурсів як джерела знань. Це забезпечує динамічне поповнення бази знань прикладних інтелектуальних систем. Наведено приклад використання запропонованого методу у задачі перетворення природномовного тексту в жестову мову в системі інформаційної підтримки осіб з вадами мовлення та слуху.

***Ключові слова**: формальна онтологія, онтологічні мови, формальна модель онтології, інтерпретація онтологій, семантичні Wiki-ресурси, інформаційна система.*

ГРИЦЕНКО В.І., ГЛАДУН А.Я., РОГУШИНА Ю.В., 2018

ВСТУП

Побудова і впровадження інтелектуальних систем, що ґрунтуються на формалізації та повторному використанні знань, є перспективним напрямом практичного застосування методів штучного інтелекту. Основу таких систем складають формалізовані подання знань про предметну область, наприклад, у формі онтології. Визначення онтології Грубером [1] як специфікації певної концептуалізації залишає відкритим питання вибору формального апарату та мовних засобів для побудови такої специфікації. Також залишаються не повністю вирішеними завдання аналізу та виявлення суті онтологічного опису предметної області, визначення властивих йому обмежень та переваг. Усе це вимагає побудови та дослідження формальних моделей для різноманітних аспектів онтологічного моделювання. Отже, актуальним є дослідження методів онтологічного моделювання з використанням формальних онтологій предметних областей (ПрО) та систем онтологічного моделювання.

Формальні онтології предметних областей зараз широко використовуються різноманітними інтелектуальними Web-застосунками для інтероперабельного подання знань. Основну роль у процесі створення моделей знань про ПрО відіграють онтології — знання, організовані на основі математичної логіки.

ПОСТАНОВКА ПРОБЛЕМИ

На відміну від динамічних Wiki-ресурсів, онтології є відносно стабільними моделями знань, однак часто не враховують змін в знаннях, особливо у нових предметних областях і тому містять застарілі дані. Оновлювати версії онтологій вручну занадто складно та дорого. Найбільш ефективним підходом є використання структури класів вже наявної онтології для семантичної розмітки Wiki-ресурсу, який може змінюватись досить часто, а потім за цією розміткою автоматизовано генерувати нові версії вказаної онтології, які можна зберігати у репозиторії онтологій.

Для того, щоб оцінити складність цієї оновлюваної онтології, тобто пов’язати її з певною конкретною логікою з сімейства дескриптивних логік, необхідно перш за все явно визначити всі характеристики формальної моделі цієї онтології. Здобути ці відомості з семантичного Wiki-ресурсу, на жаль, неможливо. Це пов’язано з тим, що, незважаючи на велику подібність онтології та семантичного Wiki, вони мають все ж таки зовсім різну виразність. Таким чином, ця частина роботи не може бути автоматизованою – якщо виникають такі зміни, то їх має вносити людина. Характеристики формальної моделі онтології можна імпортувати з попередньої версії онтології ПрО та змінювати тільки тоді, коли ці характеристики онтології дійсно змінюються. Для того, що б явно вказати, з яких джерел та які саме знання треба здобувати, ми пропонуємо розширену, детальнішу формальну модель онтології ПрО.

Метою роботи є розроблення моделей структурованого подання знань у Wiki -ресурсах на основі онтологій та методів їх застосування для вдосконалення та поповнення онтологій ПрО. Запропонований підхід надасть

змогу інтегрувати поточні відомості про зміни у предметній області та створювати актуальні онтології для різних прикладних інформаційних технологій (ІТ), що використовують онтології.

ФОРМАЛЬНА ОНТОЛОГІЯ ЯК ЗАСІБ ОПИСУ ЗНАНЬ ПРЕДМЕТНИХ ОБЛАСТЕЙ

До ІТ цей термін пришов з філософії, де формальні онтології (Ф. Брентано, Е. Гуссерль, Г. Фреге, К. Твардовський, А. Райнах, Н. Коккьярелла, С. Лесьневський, Р. Карнап, Р. Полі) відображають взаємозв'язки об'єктів (реальних та віртуальних) з їх властивостями та частинами [2–5].

Дослідження формальних онтологій у філософії — це галузь досліджень, яка із самого початку нерозривно пов'язана з логікою та її апаратом і походить ще від таких робіт Аристотеля, як «Метафізика» і «Категорії». Докладніше це проаналізовано у [6].

Термін «формальна онтологія» було вперше введено Е. Гуссерлем у його «Логічних дослідженнях» (1901), де він розрізняє формальну логіку з одного боку, та формальну онтологію — з іншого. Формальна логіка має справу з взаємозв'язками виводимості, несуперечністю і загальною значущістю певних тверджень, тоді як формальна онтологія стосується взаємозв'язків об'єктів та їх властивостей [2]. Логіка вважається формальною тому, що відношення виводимості можуть бути застосовані до логічних висновків у силу лише своєї форми, тоді як онтологія вважається формальною в тому розумінні, що оперує довільними об’єктами та їх відношеннями з об'єктами всіх матеріальних сфер чи областей реальності.

Е. Гуссерль у своїх роботах не використовує будь-який формальний апарат у сучасному розумінні змісту цього слова. Його теорія не охоплює вертикальні відношення між частинами і цілим та стосується інтуїтивних основ і понять для розроблення систем формальної онтології [2].

Однак є інший підхід до формальних онтологій, який не ґрунтується на мереології. Мереологія (від грецького «частина» і «вивчення») — формальна теорія про частини і пов’язані з ними поняття. Цей термін використано польським математиком Станіславом Лесьневським (львівсько-варшавська математична школа) [7], який аналізує філософські, логічні і математичні складники основ математики. Відношення «частина-ціле» є винятково важливим тому, що воно утворює основу поняття системи, яке часто використовують в сучасному науковому пізнанні. У мереології відношення «є частиною» (part-of) має такі властивості: 1) асиметрія; 2) транзитивність. Цю теорію розвивали К. Айдукевич, А. Тарський, Т. Котарбінський [8].

Таким чином, особливе місце серед формальних онтології займає система С. Лесьневського, яка складається з трьох компонентів — мереології, онтології та прототетики. Його онтологія задовольняє визначенню формальної онтології в розумінні Н. Коккьяреллі, у той час як його інша система — мереологія, що надбудовується над онтологією, підпадає під визначення формальної онтології Е. Гуссерля. Онтологія С. Лесьневського є як логічною системою, так і системою формальної онтології.

Н. Коккьяреллі визначає формальну онтологію як теорію буття з точки зору формальної логіки. Вважаючи, що кожна наука розглядає свій специфічний «спосіб існування», він стверджує, що формальна онтологія вивчає

різні формалізації, які належать до систематичної класифікації всіх «способів» чи категорій існування в самому загальному вигляді.

Р. Полі в «Формальній онтології» розглядає такі типи онтологій [2]: *онтологія об'єктів і властивостей,* яка описує сутність об'єктів з їх властвостями *стратифікована онтологія*, яка описує світ у категоріях залежності; *онтологія подій*, яка описує динамічні сутності; *комбінаторна онтологія*, яка описує універсум як композицію елементів і комбінацію елементів; *трансформаційна онтологія*, яка враховує внутрішні і зовнішні трансформації; *онтологія ментальних конструкцій*, яка описує ментальну активність безпосереднього споглядання та не розглядає факти зовнішнього світу.

Крім того, Р. Полі класифікує онтології за рівнем формалізації на дескриптивні, формальні та формалізовані, а за залежністю від конкретної предметної області — на локальні і глобальні онтології.

Дескриптивна онтологія має справу із сукупностями об'єктів, ґрунтується на поверхневій наочній інформації та описує ці сукупності (в деякій ПрО або глобально).

Формальна онтологія фільтрує, кодує та організує результати дескриптивної онтології (як локальної, так і глобальної). Така онтологія є формальною в розумінні Е. Гуссерля, тобто має справу з такими категоріями, як мовлення, процес, матерія, ціле, частина, число, які характеризують різні аспекти чи типи реальності, але не використовують будь-які специфічні формалізми. Формальне кодування в строгому значенні цього слова застосовують тільки на третьому рівні — на рівні побудови формалізованої онтології. Рівень формалізованих конструкцій також припускає оцінювання виразної, обчислювальної, когнітивної адекватності різних формалізмів.

Формалізована онтологія обов'язково припускає вибір формалізму, що потребує вибору певної формальної мови для опису онтології. Потреба в подібних мовах давно вже почала відчуватися розробниками систем штучного інтелекту.

Конструюванням таких формальних мов займаються Б. Сміт і К. Малліган [6], які прагнуть будувати мови на зовсім інших засадах (на відміну від того, як це робиться в логіці) та наводять опис онтологічної мови типу діаграм Венна. Але ця розробка не була орієнтована на ІТ.

ФОРМАЛЬНІ МОВИ ПОДАННЯ ОНТОЛОГІЙ ТА ЇХНЯ ВИРАЗНА ПОТУЖНІСТЬ

GOL (General Ontological Language — загальна онтологічна мова) було розроблено у 2001 р (Б. Сміт, Б. Хеллер, Х. Херре і В. Деген) для описування онтології верхнього рівня, орієнтованої на завдання медичної діагностики. На відміну від інших проектів подібного роду, GOL не є теоретико-множинною мовою [9–10]. У цьому проекті вихідні сутності реального світу поділяються на множини, індивіди та універсалії. Індивіди, у свою чергу, розділяються на моменти, субстанції, хроноїди, топоїди і ситуоїди.

KIF (Knowledge Interchange Format — формат обміну знаннями) — формальна мова для обміну знаннями між комп'ютерними програмами, написаними різними програмістами в різний час і на різних мовах. Онтологія цього проекту використовує об'єкт як найбільш широку категорію. Об'єкти можуть бути конкретними (пристрій, людина, молекула) або абстрактними (поняття

справедливості, число «два»), простими чи складними. Крім того, об’єкти поділяються на індивіди і множини. Множина є сукупністю об'єктів, індивідом є об'єкт, який не є множиною. Функції та відношення вводяться як множини скінчених списків. Онтологічний базис KIF слабкіший за базис GOL: KIF можна розглядати як теоретико-множинну підмножину GOL.

F-Logic — онтологічна мова, що базується на логіках першого порядку, однак класи і властивості в цій мові надано як терміни, а не як предикати. Мова створювалася для здійснення взаємодії між онтологіями, побудованими на основі предикатів, і онтологіями, побудованими на основі F-Logic. Такий переклад зберігає логічні зв'язки для великої кількості онтологічних мов, у тому числі і для OWL DL.

OCML (Operational Conceptual Modeling Language) — мова, розроблена для підтримки кількох типів конструкцій подання знань. Вона дозволяє задавати специфікацію й операціоналізацію функцій, зв'язків, класів, екземплярів і правил. Мова містить механізми для опису онтологій і методів рішення задач.

LOOM — мова подання знань і міркувань в галузі штучного інтелекту. Вона використовується для побудови дедуктивних виведень на основі декларативних знань.

В рамках проекту Semantic Web розроблено стандарт мови OWL. Онтологія OWL (Web Ontology Language) є послідовністю аксіом і фактів, а також посилань на інші онтології [1]. OWL-онтології містять компоненти для запису авторства та іншої необхідної інформації. OWL-онтології є документами Web, на них можна посилатися через URI (Uniform Resource Identifier). Мова OWL розширює мову RDF(S) (Resource Description Framework Schema)і забезпечує конструкції для вираження понять, відношень, потужності, анотацій та конкретизації понять тощо. Окрім того, OWL підтримується великою кількістю редакторів онтологій та різонерів. Все це дозволяє назвати OWL найбільш релевантною мовою для Semantic Web. OWL має три діалекти, що різняться за виразністю:

- OWL Lite (простота створення онтологій);
- OWL DL (повнота та можливість логічного виведення);
- OWL Full (висока виразність).

Кожен з цих діалектів (окрім Lite) є розширенням попереднього. Як наслідок, будь-яка OWL Lite онтологія є OWL DL онтологією, а будь-яка OWL DL онтологія є OWL Full онтологією.

OWL Lite — найпростіший варіант, призначений для тих користувачів, які потребують класифікації ієрархії і використовують прості обмеження. OWL Lite забезпечує швидку міграцію тезаурусів і інших таксономій.

OWL DL орієнтований на тих користувачів, які потребують максимальної виразності без втрати повноти обчислень та гарантованого завершення всіх обчислень у визначений час. OWL DL містить всі мовні конструкції OWL з обмеженнями поділу типу (клас не може бути приватною властивістю, а властивість — індивідом або класом). Терміни «приватний» або «публічний» є так званими «модифікаторами» доступу, тому що в деяких мовах вони використовуються для зміни раніше встановлених прав у спадкуванні класу. Назву OWL DL пов’язано з його відповідністю дескриптивній логіці.

OWL Full призначається для користувачів, яким потрібна максимальна виразність і синтаксична потужність RDF без обчислювальних гарантій. Наприклад, у OWL Full клас може одночасно розглядатися і як сукупність екземплярів, і як окремий екземпляр. Інша суттєва відмінність від OWL DL полягає у тому, що owl:DatatypeProperty може бути позначена як owl:InverseFunctionalProperty. OWL Full дозволяє створювати такі онтології, які розширюють склад визначеного словника RDF або OWL.

Формально описати OWL-онтологію можна таким чином:

```
<ontology> ::= Ontology ( [<authorship-etc>] {<directive>} )
<authorship-etc> ::=...
<directive> ::= <imports>
<directive> ::= <axiom>
<directive> ::= <fact>
<imports> ::= imports ( <URI> )
```

Найфундаментальніші поняття певної ПрО мають відповідати класам, що знаходяться в корені різних таксономічних дерев. Кожен екземпляр в онтології OWL належить до класу owl:Thing, а кожен встановлений користувачем клас автоматично є підкласом owl:Thing. Специфічні для ПрО кореневі класи визначаються простим оголошенням іменованого класу. OWL також задає порожній клас: owl:Nothing. Визначення можуть бути розширюваними і розподіленими.

Фундаментальним таксономічним конструктором для класів є rdfs:subClassOf. Він зв'язує окремий клас з загальним: якщо X — підклас Y, то кожен представник X — також представник Y. Відношення rdfs:subClassOf є транзитивним: якщо X — підклас Y, і Y — підклас Z, то X — підклас Z.

Визначення класу містить дві частини: назву або посилання на клас та список обмежень. Кожен вираз, який безпосередньо міститься у визначенні класу, уточнює властивості представників цього класу. Представники класу належать до перетину зазначених обмежень. Для визначення екземпляра досить оголосити його членом якогось класу.

Характеристика «властивості» дозволяє затверджувати загальні факти про членів класів і про екземпляри. Вони є бінарними відношеннями. Розрізняють два типи властивостей: *властивості-значення* — відношення між представниками класів і RDF-літералами або типами даних, зумовлених XML Schema та *властивості-об'єкти* — відношення між представниками двох класів.

Для надання властивості є багато способів обмежити це відношення. Можна задати домен і діапазон. Властивість може бути визначена як спеціалізація (підвластивість) наявної властивості. Можливі і складніші обмеження. Властивості, так само як класи, можуть бути організовані в ієрархію.

OWL використовує більшість вбудованих типів XML Schema. Посилання на ці типи здійснюються за допомоги URI для http://www.w3.org/2001 /XMLSchema. Формальна семантика OWL описує, як отримати логічні наслідки, маючи таку онтологію, тобто здобути факти, що не надано в ній безпосередньо, проте випливають з її семантики. Ці наслідки можуть бути основані на одному або множині розподілених документів, які комбінуються з використанням спеціальних механізмів OWL.

ФОРМАЛЬНІ МОДЕЛІ ОНТОЛОГІЙ

Аналіз публікацій доводить, що саме онтології є адекватним та ефективним засобом для моделювання уявлень про різні ПрО і дають змогу формально відображати їх семантику. Різні наукові джерела містять різноманітні формальні моделі подання онтологій , проте схожими рисами в них є *множина термінів* (понять, концептів), що може поділятися на множину класів і множину екземплярів, *множина відношень між поняттями*, у якій можна виділити відношення з певними фіксованими наборами властивостей — «клас-підклас», ієрархічні (таксономічні) відношення та відношення синонімії (подібності), та *аксіоми та функції інтерпретації понять і відношень*.

Будь-яка онтологія, крім множини об’єктів предметної області (класів та екземплярів), що включає характеристики об'єктів і процесів предметної області, містить опис важливих залежностей між цими об'єктами.

Формальну модель онтології можна надати у вигляді трійки

$$O = <X, R, F>,$$

де X — множина концептів, R — множина відношень між концептами, F — множина функцій інтерпретації концептів з множини X і відношень з R.

На змістовному рівні за допомоги елементів F можна вказати (задати правила визначення припустимості), по-перше, які набори концептів, їх відношень та значень їх властивостей є припустимими для ПрО — наприклад, дитина не може бути старшою за батьків, книга не може бути написана раніше дати народження її автора, препарат не може одночасно підвищувати та знижувати тиск; а по-друге, які значення властивостей та відношення між концептами можна вивести з наявної в онтології інформації – наприклад, сусід сусіда є сусідом.

Така модель має загальний характер, у той час як на практиці користуються точнішими моделями. Наприклад, у [11] онтологія визначається як структура

$$O = \langle C,\ \leq_c,\ R,\ \sigma_R,\ \leq_R,\ A,\ \sigma_A,\ T \rangle,$$

яка складається з: чотирьох множин C, R, A і T, які не перетинаються і елементи яких називають відповідно ідентифікаторами концептів, ідентифікаторами відношень, ідентифікаторами атрибутів і типами даних; структури часткового упорядкування $\leq_c$ над C з верхнім елементом $root_C$ (ієрархія концептів чи таксономією); функції $\sigma_R : R \to C^+$ (ідентифікатор, чи сигнатура відношення; часткового упорядкування $\leq_R$ над R (ієрархія відношень), де $r_1 \leq_R r_2$ означає, що $|\sigma_R(r_1)| = |\sigma_R(r_1)|$ і $\pi_i(\sigma_R(r_1)) \leq_R \pi_i(\sigma_R(r_2))$ для $\forall i : 1 \leq i \leq |\sigma_R(r_1)|$ ($\pi_i(t)$ — це і-й компонент кортежу t); функції $\sigma_A : A \to C \times T$ (ідентифікатор, чи сигнатура атрибута); безлічі типів даних T (наприклад, рядок, ціле).

Для структури $\leq$ виконуються такі умови: рефлективність, антисиметричність, транзитивність, наявність верхнього елемента і супремум.

У [12] онтологія визначається як кортеж

$$O = \langle C,\ I,\ R,\ T,\ V,\ \leq,\ \perp,\ \in,\ = \rangle,$$

в якому C — множина класів; I — множина екземплярів; R — множина відношень; T — множина типів даних; V — множина значень (множини C, I, R, T, V попарно не перетинаються); $\leq$ — відношення на $(C\times C)\cup(R\times R)\cup(T\times T)$, що називається спеціалізацією; $\perp$ — відношення на $(C\times C)\cup(R\times R)\cup(T\times T)$, що називається виключенням; $\in$ — відношення на $(I\times C)\cup(V\times T)$, яке називають реалізацією (створенням екземпляра); = відношення на $I\times P\times\cup(I\cup V)$, яку називають присвоюванням.

Така онтологія може бути перетворена у граф, вузли якого є типами.

Семантику онтологічних мов, яка описує відношення мовних виразів до позначених об'єктів та їхнього змісту [6], зазвичай подають за допомоги теорії моделей. Зокрема, вона визначає функцію інтерпретації, що відображає кожен елемент онтології на множину D — область інтерпретації.

Інтерпретацією онтології $O=\langle C,\ I,\ R,\ T,\ V,\ \leq,\ \perp,\ \in,\ =\rangle$ є пара $\langle I,D\rangle$, у якій D — область інтерпретації, а I — функція інтерпретації, така, що:

- $\forall c\in C, I(C)\subseteq D$;
- $\forall r\in R, I(r)\subseteq D\times(D\cup V)$;
- $\forall i\in I, I(i)\subseteq D$;
- $\forall t\in T, I(t)\subseteq V$;
- $\forall v\in V, I(v)\subseteq V$.

Твердження онтології задовольняє інтерпретації, якщо інтерпретація узгоджується з цим твердженням.

Узагальненою моделлю онтології $O=\langle C,\ I,\ R,\ T,\ V,\ \leq,\ \perp,\ \in,\ =\rangle$ є інтерпретація $m=\langle I,D\rangle$, яка задовольняє всім твердженням онтології о: $\forall\sigma\in o, m\models\sigma$. Зазначимо, що на практиці використовують точніші моделі з розкриттям її елементів. Для цих моделей вже можна розглядати їх інтерпретацію.

Для описання онтологій ми будемо використовувати формальну модель $O=\langle X,R,F,T\rangle$, яка складається з таких елементів:

$X=X_{cl}\cup X_{ind}$ — множина концептів онтології, де X_{cl} — множина класів, X_{ind} — множина екземплярів класів, таких, що $\forall a\in X_{ind}\exists A\in X_{cl}, a\in A$;

$R=r_{ier_cl}\cup\{r_i\}\cup r_{ier_prop}\cup\{p_j\}\cup p_{ier_prop}$ — множина відношень між елементами онтології, де r_{ier_cl} — ієрархічні відношення між класами онтології (структури часткового впорядкування з верхнім елементом Thing, що можуть встановлюватися між класами онтології і характеризується такими властивостями, як антисиметричність і транзитивність), $r_{ier_cl}: X_{cl}\to X_{cl}$; $\{r_i\}$ — множина об'єктних властивостей, що встановлюють відношення між екземплярами класів: $r_i(a, a\in X_{ind})=b, b\in X_{ind}$, $r_i: X_{ind}\to X_{ind}$; r_{ier_prop} — ієрархічні відношення між об'єктними властивостями класів онтології (структури часткового впорядкування з верхнім елементом topObjectProperty, які можуть встановлюватися між класами

онтології і властивостями класів і характеризуються такими властивостями, як антисиметричність і транзитивність), $r_{ier_prop} : \{r_i\} \to \{r_i\}$; $\{p_j\}$ — множина властивостей даних, які встановлюють відношення між екземплярами класів і значеннями з T: $p_i(a, a \in X_{ind}) = t, t \in T$, $p_i : X_{ind} \to T$; $\mathrm{p}_{\mathrm{ier_prop}}$ — ієрархічні відношення між властивостями даних екземплярів класів онтології (структури часткового впорядкування з верхнім елементом topDataProperty, які можуть встановлюватися між властивостями екземплярів класів онтології і характеризуються такими властивостями, як антисиметричність і транзитивність), $p_{ier_prop} : \{p_i\} \to \{p_i\}$;

$F = \{F_{cl} \cup F_{prop}\}$ — множина характеристик, які можна використати для логічного виведення над онтологією та задати правила ПрО: F_{cl} — множина характеристик класів онтології, що можуть застосовуватися для логічного виводу (еквівалентність, відсутність перетину тощо), F_{prop} — множина характеристик об'єктних властивостей екземплярів класів онтології, які можна застосувати для логічного виводу (транзитивність, симетричність, антисиметричність, рефлексивність, антирефлексивність тощо);

T — множина типів даних (наприклад, рядок, ціле), значення з яких можуть приймати властивості класів онтології.

ОНТОЛОГІЇ ТА ДЕСКРИПТИВНІ ЛОГІКИ

Формальна модель онтології значно спрощує співставлення її з дескриптивними логіками (DL): наявність чи відсутність певних елементів у моделі визначає складність структури цієї онтології та дозволяє робити висновки щодо складності її оброблення. DL — це сімейство мов подання знань, яке слугує для описання понять предметної області у формалізованому вигляді. Будь-яка логіка DL є логікою першого порядку, але не навпаки. Дескриптивні логіки поєднують у собі, з одного боку, досить багаті виражальні можливості, а з іншого — задовільні обчислювальні властивості, такі, як можливість розв'язання й відносно невисока обчислювальна складність основних логічних проблем, що дає змогу використовувати їх на практиці.

ALC (Attributive Language with Complements, атрибутивна мова з доповненнями) — це підмножина сімейства DL, на якій базується мова OWL; водночас для багатьох реальних онтологій цілком достатньо мови ALC, яка дає змогу описувати складні поняття за допомоги таких конструкторів класів, як перетин, об'єднання, доповнення класу, універсальне та екзистенційне обмеження властивості. Для характеристики інших DL використовують такі позначення: F — функціональні властивості; E — повна екзистенціальна кваліфікація (екзистенціальні обмеження, крім owl:Thing); U — об'єднання понять; C — складне заперечення поняття; S — логіка ALC з транзитивністю ролей; H — ієрархія ролей (підвластивості rdfs:subPropertyOf), R — обмежені складні аксіоми включення ролей; рефлексивність і антирефлексивність; диз'юнктивність ролей; O — номінали (перелічувані класи обмежень значення об'єкта owl:oneOf, owl:hasValue); I — зворотні (інверсні) властивості; N — обмеження потуж-

ності (restrictions (owl:cardinality, owl:maxCardinality)); Q — кваліфіковані обмеження потужності (наявні в OWL 2.0, обмеження потужності, відмінні від owl:Thing);(D) — використання типів даних властивостей, значень даних або типів даних.

ІНТЕРПРЕТАЦІЯ ОНТОЛОГІЙ

Для використання онтології потрібно, крім її формальної моделі, задати її семантику, тобто визначити зміст її елементів. Семантика цієї моделі може бути визначена за допомоги функції інтерпретації. Інтерпретацією онтології $O = \langle X, R, F \rangle$ є пара $\langle I, D \rangle$, у якій функція інтерпретації (I) така, що для кожного класу та екземпляра онтології є значення їх функції інтерпретації з області інтерпретації:

$$\forall x \in X, I(x) \subseteq D .$$

Те, що твердження o_k онтології O (синтаксично правильна формула відповідної онтологічної мови, яка пов'язує екземпляри або класи онтології за допомоги відношень) задовольняє інтерпретації $\langle I, D \rangle$, позначається як $I \models o_k$). Відповідно до специфіки інтерпретації онтологій O, твердження задовольняє інтерпретації $\langle I, D \rangle$, якщо виконуються умови, пов'язані з r_{ier_cl}, r_{ier_prop}, p_{ier_prop} та тими елементами множини $F = \{ F_{cl} \cup F_{prop} \}$, що використовуються в онтології.

Зберігається ієрархія класів:

$I \models r_{ier_cl}(c_1, c_2), c_1 \in X_{cl}, c_2 \in X_{cl}$ тоді й тільки тоді, коли $I \models c_1$ є підмножиною $I \models c_2$;

$I \models r_{ier_cl}(c_1, c_2), c_1 \in X_{cl}, c_2 \in X_{cl}$ тоді й тільки тоді, коли $I \models c_1$ є підкласом $I \models c_2$.

Зберігається ієрархія об'єктних властивостей класів онтології:

$I \models r_{ier_prop}(r_1, r_2), r_1 \in \{ r_i \}, r_2 \in \{ r_i \}$ тоді й тільки тоді, коли $I \models r_1$ є підмножиною $I \models r_2$;

$I \models r_{ier_prop}(r_1, \text{topObjectProperty}), r_1 \in \{ r_i \}$ тоді й тільки тоді, коли $I \models r_1$ є підмножиною $I \models$ topObjectProperty.

Зберігається ієрархія властивостей даних класів онтології:

$I \models p_{ier_prop}(p_1, p_2), p_1 \in \{ p_i \}, p_2 \in \{ p_i \}$ тоді й тільки тоді, коли $I \models p_1$ є підмножиною $I \models p_2$;

$I \models p_{ier_prop}(p_1, \text{topDataProperty}), p_1 \in \{ p_i \}$ тоді й тільки тоді, коли $I \models p_1$ є підмножиною $I \models$ topDataProperty.

Підтримується семантика еквівалентності класів онтології:

$I \models f_{eq}(c_1, c_2), c_1 \in X_{cl}, c_2 \in X_{cl}, f_{eq} \in F_{cl}$ тоді й тільки тоді, коли $I \models c_1$ дорівнює $I \models c_2$.

Підтримується семантика відсутності перетину класів онтології:

$I \models f_{\text{disjoint}}(c_1, c_2), c_1 \in X_{cl}, c_2 \in X_{cl}, f_{\text{disjoint}} \in F_{cl}$ тоді й тільки тоді, коли $I \models c_1$ не перетинається з $I \models c_2$.

Підтримується семантика симетричності об'єктних властивостей екземплярів класів онтології:

$I \models f_{\text{simmetr}}(r_1(x_1, x_2)), x_1 \in X_{ind}, x_1 \in X_{ind}, r_1 \in \{r_i\}, r_{\text{simmetr}} \in F_{prop}$ тоді й тільки тоді, коли $I \models r_1(x_1, x_2)$ співпадає з $I \models r_1(x_2, x_1)$.

Аналогічно підтримується семантика транзитивності, рефлексивності, антирефлексивності тощо об'єктних властивостей екземплярів класів онтології.

Онтологію можна розглядати як множину припустимих тверджень, побудованих з елементів формальної моделі. Інтерпретації, що не суперечать наведеним вище умовам для всіх тверджень, що входять до цієї множини, становлять можливі інтерпретації онтології: $\forall o \in O, I \models o$. Такі інтерпретації називають моделями онтології.

Виразна здатність різних формальних моделей онтологій суттєво різниться та залежить від того, які саме характеристики входять до складу $F = \{F_{cl} \cup F_{prop}\}$. Для реалізації цих моделей є мови подання онтологій, які теж різняться за своїми виразними можливостями та за своєю складністю: RDF Schemas пропонує найпростіший рівень для подання онтологій, а OWL Full — найскладніший. Слід зазначити, що чим вище виразні можливості мови, тим довше та складніше обробляти онтології, побудовані на її основі. Тому вибір засобу подання онтології залежить від специфіки завдань, для розв'язання яких вона розробляється.

Для того, щоб побудувати онтологію, потрібно визначити джерела, з яких можна здобути пертинентні знання, та розробити алгоритми, що забезпечують процес цього здобуття. Зважаючи на те, що для автоматизованого оброблення значно краще придатні структуровані дані, а для здобуття знань найбільш ефективним джерелом є семантична розмітка, розглянемо це на прикладі поповнення онтології ПрО відомостями із семантично розмічених Wiki-сторінок [13].

ПОПОВНЕННЯ ФОРМАЛЬНИХ ОНТОЛОГІЙ ПРЕДМЕТНИХ ОБЛАСТЕЙ ІЗ СЕМАНТИЧНИХ WIKI-РЕСУРСІВ

Для оновлення розробленої онтології та її розширення у разі динамічної зміни інформації про інформаційний об'єкт запропоновано метод поповнення формальних онтологій ПрО із семантичних Wiki-ресурсів (рис. 1), який об'єднує декілька етапів.

Eman 1. Створити першу версію онтології та зберегти її. Слід зазначити, якщо Wiki-ресурс уже має семантичну розмітку, яка за ПрО є близькою до онтології, що будується, то у цій онтології доцільно застосовувати визначені імена для класів, властивостей та інших елементів розмітки. Можна додавати специфічні для ПрО терміни в якості тегів розмітки.

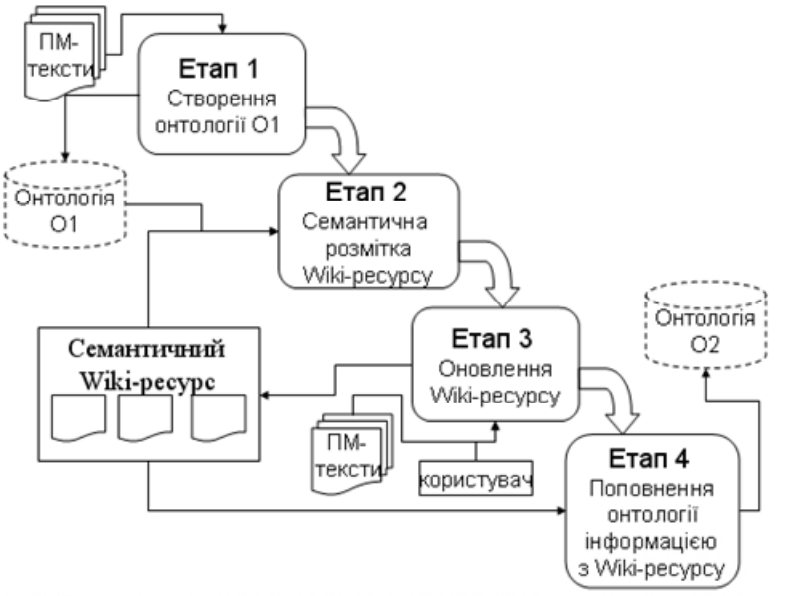

Рис. 1. Метод поповнення онтології

Наприклад, для Wiki-сайту наукового інституту ми можемо застосовувати загальновживані терміни: «Прізвище», «Ім'я», «Посада», «Назва відділу», «Публікація» та «Кількість авторів», але також додати власні терміни «Програмна платформа» та «Мета застосування агентів». Власні теги дозволять відмітити ті відомості, які належать лише онтології ПрО.

Етап 2. Розмітити семантичний Wiki-ресурс, тобто такий ресурс, який підтримує додання семантичної розмітки всіма користувачами термінами цієї онтології.

Етап 3. Дочекатися, коли у розмічені нами сторінки хтось внесе зміни — наприклад, додасть новий проект, змінить посаду тощо. Дізнатися про це ми не можемо, і тому будемо генерувати оновлення з певними інтервалами часу для перевірки.

Етап 4. Викликати оновлення зазначеної онтології, використовуючи тільки ті сторінки Wiki-ресурсу, які містять теги з цієї онтології. Відбір таких сторінок легко здійснити засобами Wiki-системи.

Розглянемо детальніше запропонований вище метод на прикладі задачі поповнення та вдосконалення онтології, призначеної для перетворення природномовного тексту з обмеженою лексикою в жестову мову.

РОЗРОБЛЕННЯ ОНТОЛОГІЧНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ТЕКСТУ ДЛЯ ЛЮДЕЙ З ВАДАМИ МОВЛЕННЯ ТА СЛУХУ

Актуальною проблемою є необхідність отримання інформації у вигляді мови жестів для людей з порушеннями мовлення та слуху. У багатьох країнах існує значний прошарок таких людей, яких потрібно активно інтегрувати у суспільство, забезпечити їх роботою, навчати. На сьогодні в Україні проживає близько 500 тисяч людей з порушеннями слуху. Для цієї категорії громадян налічу-

ється 59 спеціалізованих шкіл, 20 університетів. Є навчальні матеріали, доступні для людей з вадами мовлення та слуху (відеословники, програмне забезпечення для навчання, он-лайн курси тощо). Однак, немає ефективних засобів для автоматизованого машинного перекладу на мову та з мови жестів.

Розвиток інформаційних технологій для перекладу на жестову мову є актуальним завданням. Такая інформаційна технологія може мати великий соціальний попит, зокрема, це забезпечить людей з вадами слуху можливістю активно спілкуватися з людьми, які не володіють мовою жестів, розширити сферу спілкування і занять [14–15].

Мова жестів — це мова, яка базується на комбінації жестів. Кожен жест виконується однією або обома руками в поєднанні з мімікою та положенням тіла. Жестова мова є незалежною візуально-просторовою мовою та має свою власну граматику, яка відрізняється від граматики природної розмовної мови. Деякі функції жестової мови дозволяють паралельно переносити інформацію, виконуючи знаки обома руками, використовуючи міміку та артикуляції. В природній розмовній мові це неможливо, де передача інформації є лінійною (слово за словом).

Спільно з кафедрою прикладної математики Варшавського університету виконується соціальний проект створення інтелектуальної інформаційної системи перетворення тексту певної природної мови у жестову мову для людей з вадами мовлення та слуху, який має за мету розпізнавання природомовного тексту на основі онтології, яка допомагає ураховувати всі особливості узгодження жестової мови та природної мови [16–17]. Передбачено створення універсального проекту для багатьох європейських мов.

Для перетворення тексту у мову жестів як область інтерпретації треба одержувати відомості про синоніми, омоніми тощо різних термінів з енциклопедичних Wiki-ресурсів, що містять семантичну розмітку.

Серед завдань цього проекту — створення інтелектуальної інформаційної технології для перекладу тексту української (польської) мови у мову жестів для глухонімих людей на основі онтології, яка надає модель знань про певну предметну область. Передбачається, що літературна текстова українська (польська) мова повинна бути заздалегідь приведена до так званої контрольованої мови, тобто до подання (розбиття) складних речень через набір простих (типу — об'єкт, суб'єкт, предикат). Система повинна надати у відповідність словам контрольованої мови конкретний жест з бази даних жестів за відповідним алгоритмом (рис.2). За цих умов можливі ситуації, які потребують подальшого вирішення.

1. Є *власні імена*, які не подано жестами (наприклад, власні імена людей, географічні назви тощо). Ці імена потрібно подати за допомоги спеціального застосунку в жестову анімацію букв пальцями рук.

2. Є *фрази або речення*, які можна замінити одним жестом. Їх повинна розпізнати онтологічна інтелектуальна система.

3. Є *омоніми*, які також повинна розпізнати інтелектуальна система, наприклад, слово «коса» може мати три різних семантичних значення.

4. Є *синоніми*, які також повинна розпізнати інтелектуальна система, наприклад, ЕОМ, комп'ютер, ноутбук, мейнфрейм. Вони означають одне й те саме (відношення синонімії).

Рис. 2. Алгоритм оброблення омонімії в процесі перетворення природно-мовного тексту на жестову мову

Після оброблення (контролю) тексту інтелектуальною системою виконується послідовна інтеграція цих різнорідних компонент, їх розмітка для подальшого подання жестовою мовою.

Для удосконалення роботи інтелектуальної системи потрібно застосувати алгоритми машинного навчання, які забезпечують адаптацію системи до різних ситуацій лексичного оброблення тексту.

Зазначені інструменти можуть бути використані для розроблення системи машинного перекладу з мови розмовної в жестову і навпаки, щоб полегшити спілкування людям з вадами мовлення та слуху з тими, хто не володіє жестовою мовою.

У розробленні цього проекту вирішується низка завдань:

1) складання специфікації завдань для розроблення інтелектуальної системи розпізнавання проекту та розроблення архітектури системи;

2) розроблення специфікації завдань для філологів української (польської) мови для наповнення бази даних з розпізнавання тексту (список синонімів, антонімів, дієслово = іменник, екземплярів власних назв, груп словосполучень для одного жесту, слів, які не мають жесту, тощо);

3) розроблення концепції функціонування системи розпізнавання з урахуванням можливостей методів машинного навчання (на основі правил SWRL з використанням вхідних тестових наборів — наприклад, перетин тезаурусів семантичної оболонки для одного терміну);

4) розроблення онтологічної моделі розпізнавання тексту для виявлення в ньому неоднозначностей, багатозначностей та розмітки цього ткесту для перетворення в систему жестів;

5) розроблення методики та алгоритмів поповнення формальних онтологій розпізнавання на основі відкритих інформаційних ресурсів (Wiki-ресурсів тощо);

6) розроблення завершальних алгоритмів для програмування платформи інтелектуальної системи розпізнавання тексту та тестування цієї платформи.

Розглянуті формальні моделі онтологій та методи дозволять автоматично оновлювати версії онтологій на основі нових відомостей з відкритих ресурсів. Особливо важливим є те, що запропонованим методом значно легше розширювати словник та будувати його спеціалізовані версії для різних професійних сфер або ПрО. Використання зовнішніх баз даних забезпечує більшу повноту та зв'язність побудованого словника. Нові слова з ПрО додають автоматично, що є особливо важливим для тих нових галузей, які швидко змінюються та розвиваються, наприклад для ІТ. А саме в цій сфері працює велика кількість осіб вказаної групи. Таким чином, запропонований підхід забезпечить покращення якості життя для багатьох людей, спростивши їм спілкування у розширеній сфері. Крім того, такі системи мають велике значення у сфері освіти, для навчання людей з вадами мовлення та слуху релевантним для них спеціальностям, що має подальші перспективи розвитку цього напряму.

Інтегрована програмна технологія побудови індексів бази контекстів омонімів різних типів (функціональних, лексичних) включає модулі створення й ведення індексу омонімів, модуль узгодження індексної бази з основним лінгвістичним ресурсом — граматичним словником, а механізми виконання зовнішніх запитів дозволяють пошук типових омонімічних контекстів у текстовому корпусі на основі індексу омонімів. Всі перераховані завдання вирішуються у разі взаємодії підсистем за спеціальними запитами. База контекстів омонімів є динамічним зовнішнім ресурсом, нарощуваним в наслідок функціонування системи.

Технологія дозволу багатозначності є комплексною технологією, яка поєднує три розроблені програмні технології. Перша технологія — технологія дозволу функціональної омонімії на основі контекстних правил. Метод дозволу багатозначності на основі контекстних правил вирішує функціональну (граматичну) омонімію на основі контекстних правил, які формулюються як результат ретельної лінгвістичної експертизи поводження омоніма в сучасних мовах.

Семантично розмічені сторінки можна отримати з довільного зовнішнього семантичного Wiki-ресурсу, який є релевантним предметній області аналізу. Наприклад, це може бути сторінка енциклопедії, що відповідає терміну, який має омоніми, або ж відповідна сторінка Вікіпедії.

Використання саме семантично розмічених ресурсів має низку переваг.

- Можливість використовувати для аналізу не тільки категорії та зв'язки між сторінками, але й семантичні властивості цих зв'язків (наприклад, враховувати контент тільки тих сторінок, які пов'язані з вибраною сторінкою відношенням «складається з» або «є засобом для»).

- Можливість явно визначити, які саме зв'язки потрібно використовувати:

{{#ask:[[Категорія:ВУЕ]]
[[Категорія:Інструменти]]
[[має приклади::{{PAGENAME}}]]
| format=ul
|limit=100
}}.

- Побудова семантичних запитів для терміну, що буде автоматизовано інтегрувати всі відомості про термін на одну сторінку.

{{#ask:[[Категорія:ВУЕ]] [[Коса]]
| format=ul
|limit=20
}}

На основі семантичної розмітки зі сторінок отримуються назви термінів: виділяються посилання та семантичні посилання на інформаційні об'єкти типу «Сторінка». Наприклад:

*..."'[[Ім'я::Альберт Петров]]'", бл. [[Рік народження::1885]], [[**місце народження**: **Біла коса**]], реформатор собаководства, природознавець, філософ,...*

Для виконання завдань аналізу відношень між словами та поняттями онтології слугують так звані лексичні та лексико-семантичні бази даних або структуровані лексикони. До таких систем належать Прінстонський WordNet, MindNet, програмний продукт проекту дослідницького відділу Майкрософт, FrameNet, розроблений в університеті Берклі, VerbNet, HowNet, ConceptNet, словник української мови СУМ-11, словник Грінченка тощо. Усі сучасні лінгвістичні ресурси допомагають встановлювати значення омонімів, полісемантичних слів, виявляти синоніми, розробляти нові та оновлювати вже створені онтології чи тезауруси, автоматизовано підсумовувати тексти, розширювати можливості пошукових систем тощо [18–19]. Кожна з баз даних має як переваги, так і свої обмеження для розв'язання прикладних завдань. Серед вказаних ресурсів WordNet набув широкого застосування як найнеобхідніший словник і найуніверсальніший інструмент у галузі комп'ютерної лінгвістики. Словник СУМ-11 має сьогодні таке ж значення для української мови і сприяє сфері опрацювання природно-мовних ресурсів та лінгвістики загалом, адже тезаурус такого типу є також надійним джерелом інформації про конкретну мову.

ВЛАСТИВОСТІ ПОБУДОВАНОЇ ОНТОЛОГІЇ

Розроблена онтологія з омонімії певної мови (рис. 2):

$$O_{df} = \langle X, R, 0, 0 \rangle,$$

де із формальної моделі $O = \langle X, R, F, T \rangle$ для потреб нашого проекту взято тільки терміни та відношення між ними, F та T відсутні, дозволяє визначити відношення ієрархії та синонімії між термінами, а за контекстом встановлювати омонім, який трапляється у тексті.

Запропонована онтологія відповідає дескриптивній логіці SHOIQ (тобто SHIQ з номіналами, які дають змогу здійснювати описи класів у вигляді переліку об'єктів), що є розширенням ALC і, відповідно до описаної вище формальної моделі та її інтерпретації, може бути надана мовою OWL DL, що гарантує скінченність процесу оброблення цієї структури знань [20].

ВИСНОВОК

Дослідження математичних моделей інтелектуальних систем, які використовують онтології та моделі завдань, дають змогу визначити їхні обмеження, розвязувати завдання валідації та верифікації систем онтологічних моделей. Розроблені в роботі методи застосовано для побудови бази знань прикладних інтелектуальних систем з можливістю її динамічного поповнення.

Запропонований метод поповнення формальних онтологій предметних областей із семантичних Wiki-ресурсів забезпечує розширення словника та побудову його спеціалізованих версій для різних професійних сфер або предметних областей з використанням зовнішніх баз даних. Автоматичне додання нових слів з предметних областей є особливо важливим для тих нових галузей, які швидко змінюються та розвиваються, особливо для галузі інформаційних технологій, в якій працює велика кількість осіб з вадами мовлення та слуху. Запропонований підхід забезпечить покращення якості життя для багатьох людей, розширюючи межи їх спілкування.

Застосування інтероперабельного подання знань у вигляді онтологій та відповідних технологій і форматів дає змогу інтегрувати мультилігвістичну систему зі знаннями, наданими в Semantic Web. Розвитком запропонованого методу має стати створення глобальної семантичної мультилінгвістичної інформаційної системи для перетворення тексту на основних європейських мовах в жестову мову для людей з вадами мовлення та слуху як важливий соціальний проект.

ЛІТЕРАТУРА

1. Gruber, T.R. A translation approach to portable ontology specifications. *Knowledge Acquisition*. 1993. Vol 5. P. 199–220.
2. Poli R. Decriptive, Formal and Formalized Ontology. In: Husserl's Logical Investigations reconsidered, by Denis Fisette (ed.). Dordrecht, Kluwer Academic Publishers, 2003, 24p. — https://www.ontology.co/essays/descriptive-ontologies.pdf
3. Degen W., Heller B., Herre H., Smith B. *GOL: A General Ontological Language*, 2002, 63p. https://www.researchgate.net/publication/2498727_GOL_A_General_Ontological_Language
4. Genesereth V.R., Fikes R.E. Knowledge Interchange Format, Version 3.0, Reference Manual. Logic Group Report Logic-92-1, Computer Science Department, Stanford University. 1992. 68p. —https://pdfs.semanticscholar.org/b20c/866025b85f165557235a68143c42d53fa70f.pdf
5. Guarino N. Formal Ontology and Information Systems. In: Formal Ontology in Information Systems. Proceedings of FOIS'98, by N. Guarino (ed.). Trento. Italy, Amsterdam, IOS-Press. 1998. P. 3–15 https://klevas.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKoncepcinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf
6. Васюков В.Л. *Формальная онтология и искусственный интеллект.* М.: ИФ РАН, 2006. — 140 с.
7. Рогушина Ю.В., Гладун А.Я. Мереологические аспекты онтологического анализа интеллектуальных Web-сервисов. *Збірник праць VII Міжнародної конференції «Інтелектуальний аналіз інформації» ІАІ-2007, 12-14 травня 2007, Київ*. С. 312–321.

8. Kotarbiński T. *Elementy teorji poznania, logiki formalnej i metodologji nauk*. Lwow, 1929. — 232p.
9. Wille R., Ganter B. *Formale Begriffsanalyse*. Berlin-Heidelberg: Springer-Verlag, 1996, 192 p.
10. Ganter B., Stumme G., Wille R., eds. Formal Concept Analysis: Foundation and Application. *Lecture Notes in Artificial Intelligence*. 2005, No. 3626. 349p. — https://link.springer.com/content/pdf/bfm%3A978-3-540-31881-1%2F1.pdf
11. Lammari N., Metais E. Building and maintaining ontologies: a set of algorithms. *Data Knowledge Engineering*, 2004, No. 48 (2), P 155–176.
12. Guarino N. Formal Ontology and Information Systems. In: Ontology in Information Systems. Proceedings of FOIS-08, Trento. Italy, by N. Guarino (ed.), Amsterdam, IOS-Press, 2009. 340 p. http://mba.eci.ufmg.br/downloads/recol/FormalOntologyinInformationSystems2008.pdf
13. Рогушина Ю.В., Прийма С.М., Строкань О.В. *Створення та використання семантичних Wiki-ресурсів: навчальний довідник*. Мелітополь: ФОП Однорог Т.В., 2017. 169 с.
14. Анисимов А.В., Лиман К.С., Марченко А.А. Методы вычисления мер семантической близости слов естественного языка. *Штучний інтелект*, 2009, №3. С. 612–617. http://www.nbuv.gov.ua/portal/natural/ii/2010_3/AI_2010_3%5C3%5C00_AnisimovLiman_Marchenko.pdf
15. Lozynska O., Davydov M. Information technology for Ukrainian Sign Language translation based on ontologies. *An International quarterly journal* ECONTECHMOD, 2015, Vol. 04, No. 2, P. 13–18.
16. Гладун А.Я. Хала К.О *Стандарт ДСТУ ISO/IEC 2382:2018 Інформаційні технології. Словник*. Київ: УкрНДНЦ, 2018. 526 с.
17. Anisimov A., Marchenko O., Taranukha V., Vozniuk T. Semantic and Syntactic Model of Natural Language Based on Non-negative Matrix and Tensor Factorization. In: Proceedings of the International Conference on Natural Language Processing, 2014, Springer, Cham. P.177–184.
18. Гладун А.Я., Рогушина Ю.В. Репозиторий онтологий как средство повторного использования знаний для распознавания информационных объектов. *Онтология проектирования,* №1, 2013. С. 35–49.
19. Гладун А.Я., Рогушина Ю.В. *Семантичні технології: принципи та практики*. К: Універсаріум, 2016. 314с.
20. Гладун А.Я., Рогушина Ю.В. Основи методології формування тезаурусів з використанням онтологічного та мереологічного аналізу. *Искусственный интеллект*, 2008, №5. С. 112–124.

Отримано 03.04.2018

REFERENCES

1. Gruber, T.R. A translation approach to portable ontology specifications. *Knowledge Acquisition*. 1993. Vol 5. P. 199–220.
2. Poli R. Decriptive, Formal and Formalized Ontology. In: Husserl's Logical Investigations reconsidered, by Denis Fisette (ed.). Dordrecht: Kluwer Academic Publishers, 2003, pp. 183–210. 24p. — https://www.ontology.co/essays/descriptive-ontologies.pdf
3. Degen W., Heller B., Herre H., Smith B. *GOL: A General Ontological Language*. 2002, 63 p. https://www.researchgate.net/publication/2498727_GOL_A_General_Ontological_Language
4. Genesereth V.R., Fikes R.E. Knowledge Interchange Format, Version 3.0, Reference Manual. Logic Group Report Logic-92-1, Computer Science Department, Stanford University. 1992. 68p. — https://pdfs.semanticscholar.org/b20c/866025b85f165557235a68143c42d53fa70f.pdf
5. Guarino N. Formal Ontology and Information Systems. In: Formal Ontology in Information Systems. Proceedings of FOIS'98, by N. Guarino (ed.). Trento. Italy, Amsterdam, IOS-Press. 1998. P. 3–15 https://klevas.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKoncepcinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf
6. Vasyukov V.L. Formal ontology and artificial intelligence (monograph). Moscow: IF RAS, 2006. 140 p. (in Russian)

7. Rogushina Yu.V., Gladun A.Ya. Mereological aspects of the ontological analysis of intelligent Web-services. *Proceedings of the VII International Conference "Intellectual Analysis of Information" IAI-2007*, *12-14 Mai 2007,* Kyiv. — P. 312–321. (in Russian)
8. Kotarbiński T. Elements of the theory of cognition, formal logic and methodology of sciences. Lviv, 1929. 232p.
9. Wille R., Ganter B. *Formale Begriffsanalyse*. Berlin-Heidelberg: Springer-Verlag, 1996, 192 p.
10. Ganter B., Stumme G., Wille R., eds. Formal Concept Analysis: Foundation and Application. *Lecture Notes in Artificial Intelligence*. 2005, No. 3626. 349 p. — https://link.springer.com/content/pdf/bfm%3A978-3-540-31881-1%2F1.pdf
11. Lammari N., Metais E. Building and maintaining ontologies: a set of algorithms. *Data Knowledge Engineering*, 2004, No. 48 (2), P 155–176.
12. Guarino N. Formal Ontology and Information Systems. In: Ontology in Information Systems. Proceedings of FOIS-08, Trento. Italy, by N. Guarino (ed.), Amsterdam, IOS-Press, 2009. 340 p. http://mba.eci.ufmg.br/downloads/recol/FormalOntologyinInformationSystems2008.pdf
13. Rogushina Yu.V., Priyma S.M, Strokan O.V. Creating and using semantic Wiki-resources: tutorial. Melitopol, FOP Odnorog T.V., 2017. 169 p. (in Ukrainian).
14. Anisimov A.V., Lyman K.S., Marchenko A.A. Methods for computing measures of semantic proximity of natural language words. *Artificial Intelligence*, 2009, № 3. P. 612–617. (in Russian)
15. Lozynska O., Davydov M. Information technology for Ukrainian Sign Language translation based on ontologies. *An International quarterly journal* ECONTECHMOD, 2015, Vol. 04, No. 2, P. 13–18.
16. Gladun A.Ya. Khala K.O. Standard DSTU ISO/IEC 2382:2018 Information Technologies. Dictionary. Kyiv: UkrNDNC, 2018. 526 p. (in Ukrainian).
17. Anisimov A., Marchenko O., Taranukha V., Vozniuk T. Semantic and Syntactic Model of Natural Language Based on Non-negative Matrix and Tensor Factorization. *Proceedings of the International Conference on Natural Language Processing*, 2014, Springer, Cham. P. 177–184.
18. Gladun A.Ya., Rogushina Yu.V. Ontology Repository as a Tool for Reusing Knowledge for Recognizing Information Objects. *Ontology of Design*, No 1, 2013. P. 35–49. (in Russian).
19. Gladun A.Ya., Rogushina Yu.V. *Semantic Technologies: Principles and Practices*. Kyiv: Universarium, 2016. 387 p. (in Ukrainian).
20. Gladun A.Ya., Rogushina Yu.V. Bases of Methodology of Formation of Thesauruses with Use Ontologic and the Mereologic Analysis. *Artificial Intelligence*, 2008, № 5. P. 112–124. (in Ukrainian).

Received 03.04.2018

В.И. Грищенко[1], член-корр. НАН Украины,
директор Международного научно-учебного центра
информационных технологий
НАН Украины и МОН Украины
e-mail: vig@irtc.org.ua
А.Я. Гладун[1]**,** канд.техн.наук,
старш. науч. сотр., отд. комплексных исследований
информационных технологий и систем
e-mail: glanat@yahoo.com
Ю.В. Рогушина[2]**,** канд.физ.-мат. наук, старш. наук. сотр.,
отд. автоматизированных информационных систем
ladanandraka2010@gmail.com
[1] Международный научно-учебный центр информационных
технологий и систем НАН Украины и МОН Украины,
пр. Академика Глушкова, 40, г. Киев, 03187, Украина
[2] Институт программных систем НАН Украины,
пр. Академика Глушкова, 40, г. Киев, 03187, Украина

МОДЕЛИ И МЕТОДЫ ИСПОЛЬЗОВАНИЯ СЕМАНТИЧЕСКИХ WIKI-РЕСУРСОВ КАК ИСТОЧНИКА ЗНАНИЙ ДЛЯ ПОПОЛНЕНИЯ ФОРМАЛЬНЫХ ОНТОЛОГИЙ ПРЕДМЕТНЫХ ОБЛАСТЕЙ

Рассмотрена целесообразность применения онтологий для представления знаний в системах искусственного интеллекта, ориентированных на функционирование в открытой среде Web. Проанализированы исследования, связанные с построением формальных онтологий предметных областей, и средства их формализации. Предложена формальная модель онтологии, которая уточняет существующие подходы, более детально описывая свойства и характеристики отношений между основными элементами онтологии. Предложено использование такой модели как основы для сопоставления онтологии с дескриптивными логиками.

В работе представлен метод пополнения онтологии предметной области, который базируется на предложенной модели и использовании семантически размеченных Wiki-ресурсов как источника знаний, что обеспечивает динамическое пополнение базы знаний прикладных интеллектуальных систем. Приведен пример использования предложенного метода в задаче преобразования естественноязыкового текста в язык жестов в системе информационной поддержки лиц с недостатками речи и слуха.

Ключевые слова: *формальная онтология, онтологические языки, формальная модель онтологии, интерпретация онтологий, семантические Wiki-ресурсы, информационная система.*

V.I. Gritsenko[1], Corresponding Member of NAS of Ukraine,
Director of International Research and Training
Center for Information Technologies and Systems
of the National Academy of Sciences of Ukraine
and Ministry of Education and Science of Ukraine
e-mail: vig@irtc.org.ua
A. Ya. Gladun[1], PhD **(**Engineering**),**
Senior Researcher of the Department of Complex Research
of Information Technologies and Systems
email: glanat@yahoo.com
Yu. V. Rogushina[2]**,** PhD **(**Phys&Math)
Senior Researcher of the Department of Automated Information Systems
email: ladanandraka2010@gmail.com
[1] International Research and Training
Center for Information Technologies and Systems
of the National Academy of Sciences of Ukraine
and Ministry of Education and Science of Ukraine
40, Acad. Glushkov av., 03187, Kiev, Ukraine
[2] Institute of Program Systems
of the National Academy of Sciences of Ukraine,
40, Acad. Glushkov av., 03187, Kiev, Ukraine

MODELS AND METHODS OF THE SEMANTIC WIKI RESOURCES USE AS KNOWLEDGE SOURCES FOR RENEWAL OF FORMAL DOMEN ONTOLOGIES

Introduction. The construction and implementation of intelligent systems based on the formalization and reuse of knowledge is a promising direction for the practical application of artificial intelligence methods. The basis of such systems is formalized representations of knowledge about the subject area, for example, in the form of ontology. There remains an open question of the choice of the formal apparatus tools for the construction of ontology.

The purpose of the paper is to develop models of structured representation of knowledge in Wiki-resources on the basis of ontologies and methods of their application for improving and replenishing ontologies of the subject area. The offered approach will allow integrating the current information on changes in the subject area and creating actual ontologies for various applied information technologies using ontologies.

Results. The expediency of using ontologies for presentation of knowledge in systems of artificial intelligence oriented to functioning in the open environment of the Web is considered. The researches connected with the construction of formal ontologies of subject areas and the means of their formalization are analyzed. A formal model of ontology, which specifies the existing approaches, describing in more detail the properties and characteristics of the relations between the main elements of ontology is proposed. An example of using the proposed method in the task of transforming the natural text into a sign language in the system of information support of people with speech and hearing impairments is given.

Conclusions. The paper describes a method for renewal the ontology of a subject area based on the proposed model and the use of semantically-tagged Wiki-resources as a source of knowledge. This provides a dynamic replenishment of the knowledge base of applied intelligent systems. The proposed method of renewal formal ontologies of the subject domain from semantic Wiki-resources provides the expansion of the vocabulary and the construction of its specialized versions for various professional fields or subject areas using external databases. The automatic addition of new words from subject areas is particularly important for developing industries, especially for the IT sector, which has a large number of people with speech and hearing impairments. The proposed approach will improve the quality of life for many people, expanding the boundaries of their communication.

Keywords: *formal ontology, ontological languages, formal model of ontology, interpretation of ontologies, semantic Wiki-resources, information system.*