

УДК 681.3.06

О. Захарова

ОСНОВНІ ПРИНЦИПИ ПОБУДОВИ ОНТОЛОГІЧНОГО ГРАФ-ОРІЄНТОВАНОГО ОПИСУ ПРИКЛАДНОЇ ОБЛАСТІ

У роботі визначаються основні два етапи погодження онтологічних специфікацій: уніфікація онтологічних моделей та погодження онтологічних контекстів. Пропонується методологія побудови онтологічного опису прикладної області на базі використання граф-орієнтованої моделі представлення онтологій, відповідних правил артикуляції онтологій, алгебри онтологій та примітивних операцій перетворення графів.

Вступ

Онтологія – це явний опис концептуалізації. Онтології описують зміст, основні властивості та відносини між поняттями в базі знань. Насьогодні існує багато джерел, з якими асоційовані онтології. Онтологія може мати різноманітні форми, але обов'язково містить словник термінів, специфікацію їх сенсу, а також опис зв'язків між термінами. Тобто компоненти, з яких складаються онтології, залежать від парадигми представлення. Але всі моделі онтологій містять концепти (сутності, класи, категорії), властивості концептів (слоти, атрибути, ролі), відношення між концептами (зв'язки, залежності, функції) та додаткові обмеження (визначаються аксіомами, у деяких парадигмах фасетами).

Онтології складаються з ієрархічних описів важливих концептів у домені разом з їх описами. Ступінь формалізму може змінюватися, від природної мови до логічних формалізмів. Відповідно, можливе використання різних представлень структури онтологій.

Коли мова йде про повторне використання онтологій та інтегрування систем, що використовують різні онтології, одна з основних проблем полягає у формалізації представлення онтологій, уніфікації їх представлення, погодженні або артикуляції онтологій, що необхідні для виконання над ними дій та операцій.

Погодження онтологій

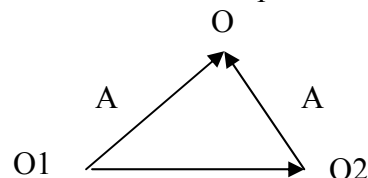
Процес погодження онтологічних специфікацій теоретично можна поділити на два етапи [1]:

- 1) уніфікація онтологічних моделей;
- 2) погодження онтологічних контекстів.

Уніфікація онтологічних моделей.

У разі роботи з неоднорідними онтологічними специфікаціями, моделі приводяться до деякого канонічного однорідного представлення, в якому будуть виконуватися всі наступні маніпуляції специфікаціями. Моделі онтологій називаються **вихідними**, а канонічна модель – **цільовою**. Задача уніфікації вихідних інформаційних моделей стає актуальною при необхідності масштабування за кількістю неоднорідних моделей. Під час уніфікації має бути побудовано відображення вихідних моделей у цільову. При відображенні моделей виконується пошук близьких конструкцій моделей та їх вираження один через одного.

При цьому доцільно звести різноманіття примітивів до обмеженого набору елементарних складових. Примітиви конкретних моделей можна виразити онтологічно за допомогою композиції понять та властивостей деякої метаонтології верхнього рівня, що містить елементарні складові, з яких можуть формуватися примітиви, що входять до онтологічної моделі. Таким чином, у специфікації онтологічної моделі O_1 та O_2 , яка описує набір її конструкцій та зв'язків між ними, елементи можуть бути анотовані описами A_{O_1} і A_{O_2} в термінах метаонтології верхнього рівня.



Такий підхід дозволяє використовувати методи семантичної інтеграції при вирішенні задачі відображення моделей.

Погодження онтологічних контекстів. Неоднорідні онтології можуть бути різними точками зору на одну й ту саму прикладну область або, що найменше, на прикладні області, що перетинаються. Ці точки зору задають онтологічні контексти, що описуються конкретними онтологіями. Тільки у цьому випадку існує сенс говорити про задачу інтеграції онтологій для погодження онтологічних контекстів.

Під **інтеграцією** у даному випадку ми розуміємо процес пошуку зв'язків між онтологіями для різних цілей. Здійснюється накопичення інформації, яка полегшує досягнення інтероперабельності між інформаційними ресурсами, що взаємодіють, та засновані на різних вихідних онтологіях. **Уніфікація** онтологій – це процес приведення онтологічних специфікацій до однорідної онтології. При цьому виконується пошук загальних частин онтологій, вирішуються конфлікти та будуються правила перетворення понять однієї онтології в іншу.

Функція перетворення однієї онтології в іншу називається **відображенням** онтологій. Відображення [2] може бути частковим у тому сенсі, що не для кожного поняття у вихідній онтології існує еквівалент в іншій онтології. Повне відображення передбачає, що інтерпретації, які мають місце для першої онтології, також мають місце й для другої. Часткове відображення означає, що в першій онтології можна виділити під онтологію, для якої існує повне відображення в другу онтологію.

Перш ніж дві онтології можуть бути відображені одна в одну, може виникнути необхідність ввести нові підтипи або супертипи понять, або відношень, щоб забезпечити відповідних кандидатів для точного відображення. Ці поняття та відношення складають **розширення** онтології у задачі уніфікації.

Глибина інтеграції онтологій може бути різною. Слабка інтеграція може використовуватися для класифікації та пошуку інформації, але не підтримувати логічні ви-

води та обчислення. Часто інтеграція вимагає більшої кількості модифікацій для підтримки більшої інтероперабельності, навіть якщо можуть залишатися де-які поняття або відношення в одній з систем, що взаємодіють, які не ідентифікуються в іншій. Взаємно-однозначна уніфікація онтологій може вимагати значних модифікацій в онтологіях.

Інший підхід – **об'єднання онтологій**. Як і відображення онтологічних понять, цей підхід включає пошук общностей та відмінностей у специфікаціях. При цьому може бути використаний метод знаходження лексичної подібності між поняттями з використанням словників та семантичної індексації. Як правило враховуються прості відношення: підклас, частина/ціле, атрибут, екземпляр, значення. Виконуються пошук зв'язків понять за їх класифікацією, пов'язуються частини, що визначають онтологію верхнього рівня, та погоджуються окремо поняття в решті частин онтологій.

«Кращий» шлях інтеграції онтологій досі є предметом інтенсивних досліджень.

Так як вихідні онтології можна змінювати, то існує відображення, що підходить (тобто, відображення без втрати інформації). Таким чином, можна в онтології додавати нові поняття та відношення, що будуть формувати аналоги, що підходять для відображення. Специфікація погодження називається **артикуляцією**. Погодження, як і відображення, може бути частковим.

Відомі різні підходи до формального представлення онтологій та роботи з ними. Так, у [3] пропонується категоріально-теоретична модель, що представляє онтології у вигляді об'єктів та працює з ними як з об'єктами. У роботі [4] пропонується підхід до представлення онтологій у вигляді кінцевого автомата та вводяться операції на онтологіях, що використовують таку модель представлення. Методологія побудови онтологічного опису, що пропонується, базується на граф-орієнтованій моделі представлення онтологій.

Граф-орієнтована модель представлення онтологій

У граф-орієнтованій моделі структура онтології представляється у вигляді зв'язаного ациклічного графа, де вузлами є терміни або концепти, а дугами – відношення між ними вказаних типів. Зв'язаний ациклічний граф, у свою чергу, може бути представлений множиною дуг, які представляються парами суміжних вузлів. Отже, він може бути інтерпретований як реляційна база даних, де ключовими колонками є пари вузлів, додатково можуть бути зазначені мітки вузлів та дуг (наприклад, назви рівнів ієрархії та видів відношень).

У загальному випадку, кожна з проблемно-незалежних моделей онтологій O може бути представлена направленим розміченим графом $G = \langle N, E \rangle$, де N – кінцева множина розмічених вузлів, а E – кінцева множина розмічених дуг [5]. Вузли графа відповідають концептам онтології, дуги графа – відношенням між ними. З кожним вузлом графа пов'язаний концепт та множина його властивостей. Дуга графа представляється трійкою (n_1, α, n_2) , де n_1 та n_2 належать до множини N (тобто є вузлами графа), α – мітка дуги. Мітка вузла n задається функцією $\lambda(n)$, яка ставить у відповідність до кожного вузла рядок, що не є вільним. У контексті онтологій, ця мітка часто є іменником, що представляє концепт. Мітка α дуги $e = (n_1, \alpha, n_2)$ є рядком, що задається функцією $\alpha = \delta(e)$. Рядок, який приєднаний до дуги, є або дієсловом природної мови, або наперед визначеним семантичним відношенням. Мітка дуги може визначати окрім імені відношення ще його кардинальність. Кардинальність відношення може задаватися або невід'ємним цілим числом, або у вигляді інтервалу на множині невід'ємних цілих чисел. Доменом функцій λ та δ є універсальна множина всіх вузлів (всіх графів), а областю значень – множина рядків (із всіх словників).

Модель графа. Для роботи з онтологічними графами необхідно визначити у скороченому вигляді частини графів, які нас цікавлять. Для цього можна використовувати моделі графів. Будемо вважати, що

модель P , яка є графом $P = (N', E')$, відповідає підграфу, якщо окрім того, що збігаються їх структури, мітки відповідних вузлів та дуг ідентичні.

Формально, вважають, що граф $G_1 = (N_1, E_1)$ відображається у $G_2 = (N_2, E_2)$, якщо існує загальна функція відображення $f: N_1 \rightarrow N_2$ така, що:

$$\forall n_1 \in N_1, \lambda_1(n_1) = \lambda_2(f(n_1)) \quad (1)$$

$$\forall e_1 = (n_1, \alpha, n_2) \in E_1, \exists e_2 = (f(n_1), \alpha, f(n_2)) \in E_2. \quad (2)$$

На практиці, окрім суворого збігу, експерт, який відповідає за сумісність домену, може визначити версії нечіткого збігу. Наприклад, експерт може вказати множину синонімів та задати правило, яке послаблює першу умову, й тим самим зробити можливою відповідність вузлів не лише у випадку точного збігу їх міток, а й у випадку, коли вони є синонімами, що задані експертом. Альтернативно, друга умова, яка вимагає збігання міток дуг, теж може дотримуватися не точно.

Перетворення графів. Для перетворення онтологічних графів використовуються чотири примітивні операції: вставка вузлів, видалення вузлів, вставка та видалення дуг. Вставка вузлів та дуг використовується під час генерації артикуляції. Видалення – операція необхідна при оновленні артикуляції, у відповідності до змін в онтологіях, що покладені в основу. Контекстом операцій є граф $G = (M, E)$, де M – множина вузлів m_1, m_2, \dots, m_n та підстрочних символів $i, j, k \in \{1, \dots, n\}$. Якщо коротко, то операції визначаються наступним чином.

1. **Вставка вузла (NA).** Задано граф G та вузол N із суміжними дугами $\{(N, \alpha_i, m_j)\}$, що необхідно додати. Результатом операції NA є граф $G' = (M', E')$, де $M' = M \cup N$, а $E' = E \cup \{(N, \alpha_i, m_j)\}$.

2. **Видалення вузла (ND).** Нехай $N \in M$ – вузол, який необхідно видалити, та $Z = \{(N, \alpha_i, m_j) \cup (m_j, \alpha_j, N)\}$ – дуги, які прилегли до N . Тоді результатом операції видалення вузла ND , на графі G , є граф $G' = (M', E')$, де $M' = M - N$ та $E' = E - Z$.

3. **Вставка дуги (EA).** Задано граф та множина його дуг $SE = \{(m_i, \alpha_j, m_k)\}$, які

необхідно додати до графа. Результатом операції $EA[G, SE]$ є граф $G' = (M, E')$, де $E' = E \cup SE$.

4. Видалення дуги (ED). Дано граф і множина дуг $SE = \{(mi, aj, mk)\}$, які необхідно видалити. Результатом операції $ED[G, SE]$ є граф $G' = (M, E')$, де $E' = E - SE$.

З метою спрощення в подальшому будемо використовувати мітки вузлів замість самих вузлів. Тобто до дуги $e = \{(n1, a, n2)\}$, де $\lambda(n1) = A$, а $\lambda(n2) = B$, будемо звертатися як $e = \{(A, a, B)\}$. Це, звичайно, не є проблемою погодження онтологій, де поняття (що представляє концепт) зображується одним вузлом в онтологічному графі. Просто поняття (мітка) та вузол графа взаємозамінні.

Для погодження онтологій у даному випадку використовується артикуляційна онтологія, що представляється артикуляційним графом.

Артикуляція онтологій. Окремі онтології, у подальшому, будемо називати **вихідними**.

Правила артикуляції визначають, які концепти у вихідних онтологіях, по одинці або у поєднанні, пов'язані між собою. Артикуляційна онтологія містить ці концепти та відношення між ними. Термін **артикуляція** буде охоплювати як артикуляційну онтологію так і правила, що задають відношення між концептами артикуляційної та вихідних онтологій.

Проблему артикуляції можна звести до двох задач. Перша – визначення семантично релевантних класів для включення їх в онтологію артикуляції, друга – створення семантичних зв'язків, тобто відношень «бути підмножиною» між класами онтології артикуляції та класами онтологій, що з ними пов'язані та покладені в основу.

Артикуляційний граф OA будується із структур, що приймаються із джерел, що покладені в його основу, та підтримує інформацію, яка стосується відносин, які існують між ними. Артикуляційний граф OA будується із використанням загальних правил артикуляції, які мають вигляд $P \Rightarrow Q$, де P та Q – складні моделі графів. Конструкція $P \Rightarrow Q$ читається як «об'єкт Q

семантично належить класу P » або « P семантично має на увазі Q ». Ця конструкція фактично представляє відношення «направлена підмножина».

Під **семантичною імплікацією** будемо розуміти правило $O1.A \Rightarrow O2.B$, де A та B прості ідентифікатори вузлів, що фактично задає одну дугу $(A, SIBridge, B)$ між структурами онтологій. Ця дуга моделює варіант, коли об'єкт A є семантичною спеціалізацією об'єкта B , та є найпростішою семантичною зв'язкою (префіксом є ім'я онтології).

Правила не обмежені описом пов'язаних онтологій, а використовуються також для структурування онтологій окремих джерел або самого артикуляційного онтологічного графа. Іншими словами, мультипонятійна імплікація (слідування) розбивається механізмом успадкування на множину атомних правил слідування.

Операнди для семантичної імплікації можна узагальнити таким чином, щоб вони охоплювали предикати моделі графа. Щоб ввести вузол, що представляє відповідний підклас, виконується трансляція предикатів на рівень даних і це приймається як мета імплікації. За замовченням, міткою для вузла N є текст предиката, який користувач може замінити ім'ям, що є точнішим та краще підходить для семантичного класу, що залучений.

Так як механізми логічного виводу для систем першого порядку, як правило, не масштабуються до великих баз знань, то більшість прикладних систем, з метою підвищення продуктивності, для представлення правил артикуляції використовують прості Класи Хорна.

Функціональні правила. Різні онтології найчастіше можуть містити терміни, які представляють одні й ті самі поняття, але виражені різними метриками простору. Тому необхідно використовувати функції нормалізації, які приймають множину вхідних параметрів, та виконують над ними бажані перетворення. Такі функції мають надаватися експертами та можуть бути записані за допомогою звичайних мов

програмування. Такі функції можна генерувати напівавтоматично.

Наприклад, ціна автомобіля може бути в різних онтологіях представлена в різних грошових одиницях, і може потребуватися її нормалізація, яка заключається в переведенні ціни у Євро, до того моменту, як ці онтології будуть інтегровані. Вибір нормалізаційної валюти має здійснюватися експертом.

Структура артикуляційної онтології. Побудова артикуляційної онтології, як вище описано, в основному зводиться до додавання вузлів та дуг між цими вузлами та вузлами вихідних онтологій. Якщо правила були сформульовані правильно, то кількість дуг буде невеликою. Такі правила імплікації особливо корисні, якщо експерт припускає створення нової структури для онтології. Важливо відмітити, що артикуляційна онтологія, яка складена з двох онтологій, може бути з'єднана з іншою вихідною онтологією, для створення нової артикуляційної онтології, яка покриває три вихідні. Це означає, що при додаванні нового джерела немає необхідності реструктурувати існуючі онтології артикуляції, а можливо їх використувати та створити нову артикуляцію з найменшими зусиллями.

Алгебра онтологій. Вона визначається [5] таким чином, щоб забезпечити взаємодію між онтологіями за допомогою артикуляційної онтології. Оператори алгебри працюватимуть з онтологічними графами. Унарні оператори (*filter* та *extract*) працюватимуть на окремій онтології. Вони є аналогічними операціям **select** та **project** в реляційній алгебрі. Ці операції допоможуть визначити області онтології, які нас цікавлять, та які будуть використовуватися в подальшому. **Унарна** операція повертає частину графа, який представляє онтологію. **Бінарні** операції – об'єднання, перетин та різниця. Кожна операція визначається на двох онтологіях (та артикуляційних правилах), а результатом буде онтологія, яка в подальшому може компонуватися з іншими онтологіями.

Вхідними для оператора **об'єднання** є два графи онтологій та набір правил арти-

куляції. В результаті виконання оператора генерується граф уніфікованої онтології, де отримана в результаті уніфікована онтологія складається з двох графів вихідних онтологій, що зв'язані артикуляцією. Артикуляція генерується за допомогою правил артикуляції. Об'єднана онтологія OU первинних онтологій $O1$ та $O2$ визначається як $O1 \cup_{rules} O2 = OU$, де *rules* звертається до множини правил артикуляції. Нехай $O1 = (N1; E1)$; $O2 = (N2; E2)$ – графи, що представляють вихідні онтології. Та нехай $OA = (NA; EA)$ представляє артикуляційну онтологію та *BridgeEdges* множина дуг, які з'єднують вузли графа OA , або $O1$ або $O2$, згідно правил артикуляції. Нехай $OU = (N; E)$ – граф, що представляє уніфіковану онтологію. $OU = (N; E)$ такий, що $N = N1 \cup N2 \cup NA$ та $E = E1 \cup E2 \cup EA \cup BridgeEdges$.

Оператор **перетину** приймає на вході два онтологічних графа, множину правил артикуляції, та видає як результат граф артикуляційної онтології. Останній складається з вузлів, що додаються згідно правил артикуляції, та дуг між цими вузлами. Щоб бути впевненим, що дуги в артикуляційній онтології дійсно з'єднують вузли, які належать графу артикуляційної онтології, дуги між вузлами в графі артикуляційної онтології та вузли з графів вихідних онтологій не включаються до артикуляційної онтології, якщо вузли у графах вихідних онтологій не є частиною графа артикуляційної онтології. Таким чином, перетин створює онтологію, що може в подальшому бути складена з іншими онтологіями. Ця операція є центральною у поняттях артикуляції. Онтологія OI перетину вихідних онтологій $O1$ та $O2$ визначається як $O1 \cap_{rules} O2 = OI$, де $OI = OA$.

Різниця двох онтологій ($O1 - O2$) визначається як поняття та відношення першої онтології, які не існують у другій. Ця операція дозволяє визначити розширення онтології, що залишається незалежним від сполучення з іншими онтологіями. Якщо зміни у вихідній онтології, наприклад, $O1$ мають місце у частині, що є різницею $O1$ з іншими онтологіями, тоді в артикуляцій-

них онтологіях ніякі зміни не мають місця. З іншого боку, якщо вузол належить $O1$, але не належить $O1 - O2$, тоді будь-які зміни, що пов'язані з цим вузлом або суміжними дугами, повинні мати місце і в артикуляційних онтологіях.

Нехай $O1 = (N1; E1)$; $O2 = (N2; E2)$ графи, що представляють вихідні онтології, $OA = (NA; EA)$ – граф, який представляє артикуляційну онтологію. Різність, $OD = O1 - O2$ представляється графом $(N; E)$ таким, що $n \in N$ тільки, якщо

- 1) $n \in N1$ та $n \notin N2$ та
- 2) не існує жодного шляху з n у будь-яке n' в $N2$ та $e \in E$ тільки в тому випадку, якщо
 - а) $e \in E1$ та
 - б) $e = (n1; \alpha; n2)$, $n1 \in N$ та $n2 \in N$.

На базі вищенаведеної структури представлення онтологій та алгебри онтологій далі буде визначена методика побудови онтологічного опису прикладної області із використанням існуючих онтологій. Для цього також будуть введені деякі додаткові операції алгебри онтологій.

Методологія побудови онтологічного опису прикладної області

Формалізація онтологій за допомогою графів вирішує проблему їх структурної різnorodності.

На онтологічному (семантичному) рівні різnorodність народжується у наслідок відмінності [1]:

- 1) імен понять та відношень;
- 2) підходів до визначення понять;
- 3) розбиття предметної області на поняття;
- 4) покриття предметної області;
- 5) поглядів на предметну область.

Таким чином, нам необхідно вирішити проблеми різnorodності онтологій на семантичному рівні, прийнявши на основу граф-орієнтовану модель представлення онтологій. Тому, першим кроком буде.

1. Уніфікація базового словника, виявлення та уніфікація еквівалентних концептів. Слід зауважити, що одні й ті самі поняття та відношення можуть мати різ-

ні імена в різних онтологічних моделях, але й концепти з однаковими іменами можуть мати різну семантику. Тому, першим кроком є синтаксичний та семантичний аналіз вихідних онтологічних моделей та визначення відображення між концептами та відношеннями (а відповідно між вузлами та дугами вихідних онтологічних графів) на базі відношень подібності або тотожності, та створення уніфікованого базового словника даного онтологічного простору.

Під **онтологічним простором** будемо розуміти об'єднану множину всіх концептів та відношень множини онтологій за виключенням повторень тотожних або подібних елементів.

Базовий словник має містити всі концепти, що належать більш ніж одній онтології з множини, та пов'язані з ними відношення, що виключає можливість виникнення протиріч. Всі інші концепти та відношення можуть бути використані в первинному вигляді, тобто, як вони були визначені у вихідній онтології, або змінені за бажанням розробника.

Таким чином, нехай існують N онтологій $O1, O2, \dots, ON$, що описують проблемно незалежні або проблемно залежні частини предметної області, для якої необхідно створити онтологічний опис. Кожна онтологія задається множиною концептів, їх властивостей та відношень між ними. Будемо вважати концепти аналогічними, якщо вони мають однакову семантику. Слід зазначити, що вони можуть мати при цьому різний синтаксис. Такі концепти заносяться до базового словника об'єднаного онтологічного простору. Кожний елемент словника можна описати набором $(NameO, EO, NameO1, NameO2, \dots, NameON)$, де $NameO$ – це уніфіковане ім'я елемента, $EO \in \{concept, relation, property\}$, $NameOi$ – ім'я елемента в онтології Oi . Якщо елемент відсутній в онтології Oi , то $NameOi = 0$.

2. Виявлення концептів, що належать різним онтологіям, пов'язані між собою, але не є еквівалентними. Якщо можна встановити де-яке відношення Rel між концептом cki , що представлений вуз-

лом Nki в онтологічному графі Oi , та концептом Cj , що представлений вузлом Nlj онтологічного графа Oj , і це відношення має певну значиму семантику в предметній області, що описується, то це відношення може бути представлене відповідною дугою (Nki, α, Nlj) в об'єднаному онтологічному графі предметної області, де $\alpha = Rel$.

Назвемо такі вузли **зовнішньо пов'язаними**.

Виявити такі вузли дозволяє також семантичний аналіз вихідних онтологій. Вузли, що представляють тотожні концепти різних онтологій, також є зовнішньо пов'язаними, так як їх можна пов'язати відношенням еквівалентності. Але в результативному графі це відношення можна опустити, а відповідні (після виконання уніфікації словника вже однойменні) вузли злити. При цьому суміжні дуги цих двох вузлів можна об'єднати з видаленням повторень.

3. Виявлення концептів перетину між онтологіями. Дві онтології Oi та Oj мають перетин, якщо вони містять тотожні концепти (після уніфікації базового словника однойменні). Відповідні онтологічні графи мають еквівалентні вузли, але може бути різною кількістю та склад суміжних дуг. Всі концепти, що належать більш ніж одній онтології, будемо іменувати концептами перетину. Кожному такому концепту буде відповідати лише один вузол в результативному онтологічному графі предметної області.

4. Виявлення концептів, що є зв'язуючими між концептами перетину, у кожній з онтологій. Нехай Ci та Cj два концепта онтології O , що представлені відповідно вузлами Nki та Nlj онтологічного графа. Зв'язуючими будемо вважати всі концепти онтології, що представлені вузлами $Nk-1i, Nk-2i, \dots, N0, N1j, \dots, Nl-1j$, де $Nk-1i$ – батьківський вузол вузла Nki в онтологічному графі, $Nk-2i$ – батьківський вузол вузла $Nk-1i$, і т. і., $N0$ – кореневий вузол графа, $Nl-1j$ – батьківський вузол вузла Nlj і т. п. Будемо називати такі концепти – **внутрішньо пов'язаними**.

5. Побудова онтологічного ядра, як мінімального онтологічного графа предметної області. Граф онтологічного ядра буде відображати основні зв'язки між вихідними онтологіями. Тому, онтологічне ядро буде містити:

- концепти перетину без повторень;
- концепти з різних онтологій, між якими доцільно встановити відношення, що будуть семантично суттєвими, без повторень;

- концепти, що є зв'язуючими між концептами перетину у кожній з вихідних онтологій, без повторень;

- відношення між цими концептами без повторень.

Відповідно, граф онтологічного ядра буде складатися з:

- вузлів, що відповідають концептам перетину у вихідних онтологічних графах, з іменами згідно базового уніфікованого словника, та за виключенням повторень. Тобто результативний онтологічний граф має містити рівно один вузол для кожного елемента базового словника;

- зовнішньо пов'язані вузли без повторень;

- внутрішньо пов'язані вузли без повторень;

- зв'язки між цими вузлами, що існують у вихідних онтологіях, з урахуванням базового словника, без повторень;

- зв'язки між зовнішньо пов'язаними вузлами;

- кореневий вузол графа, що пов'язаний з кожним з корневих вузлів онтологічних графів вихідних онтологій (всі кореневі вузли онтологічних графів вихідних онтологій належать результативному графу, за визначенням внутрішньо пов'язаних вузлів).

6. Виділення додаткових, семантично суттєвих, підграфів вихідних онтологій. Вихідні онтології можуть містити велику кількість надлишкових знань, та не містити деякі з необхідних. Тому може виявитися недоцільним повне об'єднання вихідних онтологій. З цих міркувань ми будемо додавати до результативного онтологічного

ного опису лише семантично суттєві для нас концепти та відношення.

Граф онтологічного ядра, в свою чергу, відображає лише основні зв'язки між вихідними онтологіями, та не відображає повної семантики прикладної області. Тому необхідно доповнити його семантично суттєвими концептами, що опиисуються онтологічними графами вихідних онтологій, але не потрапили до онтологічного ядра. Для цього необхідно виділити такі підграфи в онтологічних графах вихідних онтологій, визначити їх зв'язки з вузлами графа онтологічного ядра, та додати їх до графа.

7. Доповнення онтології новими концептами та відношеннями, а онтологічного графа новими вузлами та зв'язками. Для цього необхідно:

- визначити неохоплені проблемні частини предметної області, що опиисується. Будемо далі іменувати їх **семантичними додатками**;

- визначити концепти, що будуть описувати кожний з семантичних додатків, та відношення між цими концептами;

- для кожного з семантичних додатків побудувати онтологічний підграф, в якому вузли будуть відповідати визначеним концептам, а дуги відношенням між ними;

- визначити відношення між концептами семантичних додатків та концептами вихідних онтологій, та відповідно визначити зв'язки між кореневими вузлами підграфів семантичних додатків та вузлами графа, що був отриманий в п.6. Якщо додаткові підграфи були виділені правильно, то всі такі вузли, що відповідають концептам вихідних онтологій, які пов'язані з концептами семантичних додатків, будуть присутніми у результативному графі. Якщо деякий з вузлів виявиться відсутнім необхідно повернутися до п. 6.

Інший підхід до формування графа об'єднаного онтологічного простору. Інший підхід полягає в тому, щоб видалити з вихідних онтологічних графів надлишкові вузли, що відповідають концептам, які нас семантично не цікавлять, перед їх об'єд-

нанням. Після чого об'єднати отримані графи з урахуванням уніфікованого базового словника та додаткових відношень між концептами вихідних онтологій. Для цього, необхідно виконати:

- семантичний аналіз вихідних онтологій та видалити з кожного з онтологічних графів всі вузли, що не цікавлять нас у ракурсі предметної області, що опиисується, та суміжні дуги. В результаті отримаємо послідовність скорочених онтологічних графів $O1', O2', \dots, ON'$;

- синтаксичний та семантичний аналіз скорочених онтологічних графів та формування базового уніфікованого словника онтологічного простору на базі онтологічних графів $O1', O2', \dots, ON'$;

- виявлення зовнішньо пов'язаних вузлів онтологічних графів $O1', O2', \dots, ON'$;

- об'єднання онтологічних графів $O1', O2', \dots, ON'$ без повторень. Тобто всі вузли, що представлені в уніфікованому базовому словнику додаються до результативного графа однократно з ім'ям $NameO$. Тобто можна сказати, що має місце відображення $f(NameO_i) = NameO$ для всіх вузлів, що відповідають концептам, які занесені до базового словника. Таке саме відображення має місце для всіх властивостей та відношень, що представлені в базовому словнику. Вставка решти зв'язків виконується без змін. Решта вузлів всіх скорочених онтологічних графів $O1', O2', \dots, ON'$ та суміжні з ними дуги додаються у результуючий граф без змін. З кореневого вузла результуючого онтологічного графа буде виходити N дуг до кожного з корневих вузлів скорочених вихідних онтологій $O1', O2', \dots, ON'$;

- вставку дуг, що будуть відображати зв'язки між зовнішньо пов'язаними вузлами онтологічних графів $O1', O2', \dots, ON'$;

- якщо існують деякі аспекти предметної області, що залишилися поза множиною концептів, відношень та властивостей, представлених отриманим онтологічним графом, тобто що не були охоплені жодною з вихідних онтологій, необхідно:

- 1) визначити такі неохоплені проблемні частини предметної області (семан-

тичні додатки, які фактично складають розширення онтології);

2) визначити концепти, що будуть описувати кожний з семантичних додатків, відношення між цими концептами та їх властивості та для кожного з таких семантичних додатків побудувати онтологічний підграф, в якому вузли будуть відповідати визначеним концептам, а дуги відношенням між ними;

3) визначити відношення між концептами семантичних додатків та концептами вихідних онтологій, та відповідно визначити зв'язки між кореневими вузлами підграфів семантичних додатків та вузлами отриманого графа, тобто вставити підграфи семантичних додатків до результативного графа.

1. *Скворцов Н.А.* Вопросы согласования неоднородных онтологических моделей и онтологических контекстов. – Симпозиум «Онтологическое моделирование». – М.: ИПИ РАН, 2008, http://synthesis.ipi.ac.ru/synthesis/publications/08onto-hetont/_08onto_hetont.pdf
2. <http://www.obitko.com/tutorials/ontologies-semantic-web/operations-on-ontologies.html>
3. *Antoine Zimmermann, Markus Krotzsch, Jérôme Euzenat and Pascal Hitzler.* Formalizing Ontology Alignment and its Operations with Category Theory. – The Fourth International Conference on Formal Ontology in Information Systems, 2006, http://korrekt.org/page/Formalizing_Ontology_Alignment_and_its_Operations_with_Category_Theory
4. *Крытый С.Л., Ходзинский А.И.* Автоматное представление онтологий и операции на онтологиях. Институт кибернетики им. В.М. Глушкова НАН Украины, 2008, <http://shcherbak.net/avtomatnoe-predstavlenie-ontologij-i-operacii-na-ontologiyax/>.
5. *Prasenjit Mitra, Gio Wiederhold, Martin Kersten.* A Graph-Oriented Model for Articulation of Ontology Interdependencies. - Technical Report, 1999, <http://www.dit.unitn.it~p2p/RelatedWork/Matching/graphmodel.pdf>

Про автора:

Захарова Ольга Вікторівна,
кандидат технічних наук,
науковий співробітник.

Місце роботи автора:

Інститут програмних систем
НАН України,
Проспект Академіка Глушкова, 40.
Тел.: 526 5139,
e-mail: ozakharova68@gmail.com.

Отримано: 07.04.2010