

МАТЕМАТИЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ІНТЕГРАЦІЇ ДАНИХ ІНФОРМАЦІЙНИХ СИСТЕМ

Щур В.В.¹⁾, Тимчишин В.С.²⁾, Ковбасістий А.В.³⁾, Лопусевич В.П.⁴⁾

Тернопільський національний економічний університет
^{1,4)}магістрант, ^{2,3)}аспірант

I. Постановка проблеми

На сьогоднішній день актуальним завданням для організацій і установ є побудова інтегрованих автоматизованих інформаційних систем, що забезпечують підтримку різних бізнес-процесів. Отримання агрегованої інформації з метою прийняття управлінських рішень залежить від ефективної взаємодії інформаційних систем. Бізнес-процеси часто автоматизуються за використанням програмних рішень власних і сторонніх розробників без урахування їх взаємозв'язку, що особливо характерно для університетів [1].

При постійних змінах в цих бізнес-процесах розробники змушені займатися коригуванням програм і моделей даних, що призводить до структурної і семантичної неоднорідності інформаційних елементів і, відповідно, необхідність повторної розробки програмних конвертерів даних. Так як дані в підсистемах є інформаційні об'єкти, що моделюють деякі фрагменти предметної області, яким відповідає своя понятійна система (онтологія), то для вирішення проблеми семантичної неоднорідності даних необхідно узгодити розуміння предметних областей взаємодіючих об'єктів.

II. Мета роботи

Метою дослідження є розробка математичної моделі та програмної системи інтеграції даних, заснованих на застосуванні онтологічного підходу.

III. Математична модель інтеграції даних інформаційних систем

Пропонується алгоритм інтеграції на основі результатів порівняння концептів, їх атрибутів і відносин між концептами [1]. Крім того, пропонується узагальнення завдання об'єднання онтологій до процедури інтеграції онтологій, що описують різні предметні області одного і того ж домена предметних областей. Завдання інтеграції зводиться до задачі побудови відображень та інтеграції онтологій, а потім і встановлення взаємозв'язків схем інтегрованих систем, тобто збереження відповідності множини онтологій заданому набору семантичних залежностей, дозволяючи встановити взаємодію між інформаційними системами. Для опису онтологічних специфікацій використовуються RDF-твердження [2,3].

Як правило, об'єктна схема включає в себе елементи, які відповідають сутностям різних предметних областей, кожен об'єкт характеризується значеннями набору атрибутів і представляється як множина впорядкованих пар виду

$$u = \{ \langle a_i, d_i \rangle \}, \quad (1)$$

де a_i - атрибут об'єкта; d_i - значення атрибута $i \in [1..n]$; n - кількість атрибутів.

Базовим поняттям пропонованої моделі є концепт C (клас об'єктів). Кожен концепт онтології інформаційної системи визначається як одиниця знання та ідентифікується по імені і характеризується типом. Тому концепт задамо як

$$C_i = (Name_i, type_i), \quad (2)$$

де $Name_i$ - унікальне ім'я (ідентифікатор) i -го концепту; $type_i$ - тип i -го концепту (абстрактний, представлення, або складений).

Онтології інтегрованих інформаційних систем спочатку ніяк не пов'язані, отже, необхідно знайти семантично близькі елементи онтологій. Побудова математичної моделі інтеграції системи з урахуванням зіставлення їх онтологічних специфікацій створює можливість для вимірювання близькості (подібності) концептів онтології.

Метод обчислення семантичної близькості концептів дозволяє кількісно оцінити схожість між поняттями. Для кожного концепту однієї онтології формується множина релевантних семантичних концептів іншій онтології. З метою ранжирування елементів результуючої множини необхідно визначити порогові значення міри близькості. Розроблено метод класифікації рівнів близькості концептів для встановлення їх коректного відображення (рисунок 1).

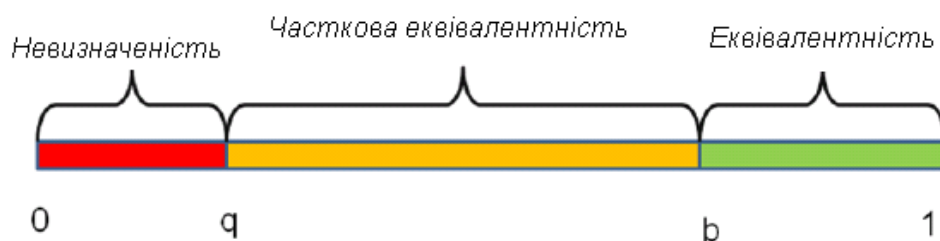


Рисунок 1 – Метод класифікації рівнів семантичного подібності концептів

Таким чином, можна побудувати модель єдиного інтегрованого інформаційного простору на основі онтологій інформаційних систем різних предметних областей, яка є уніфікованою точкою входу інформації з систем і джерел даних в єдиний інформаційний простір.

На основі запропонованої моделі розроблена програмна система інтеграції даних інформаційних систем. Програмна система дозволяє автоматично визначати і встановлювати семантичні зв'язки двох онтологій з подальшим встановленням взаємозв'язку схем інтегрованих систем, а також виконувати алгоритм формування запитів на додавання інформації з однієї системи в іншу. На основі зв'язку онтологічних понять зв'язуються релевантні елементи об'єктних схем.

У структуру програмної системи входять наступні компоненти:

- модуль створення і зміни онтологічних описів, який відповідає за створення і зміну понять і фундаментальної онтології (рисунок 2);
- репозиторій метаданих, що забезпечує зберігання опису онтологій предметних областей;
- модуль зіставлення та інтеграції онтологій, що виконує злиття онтологій різних предметних областей;
- модуль вирішення конфліктів, що виникає при зіставленні концептів онтологій;
- модуль формування шаблону запитів на додавання інформації. На основі зв'язку онтологій створюється SQL-запит на додавання інформації;
- призначений для користувача інтерфейс.

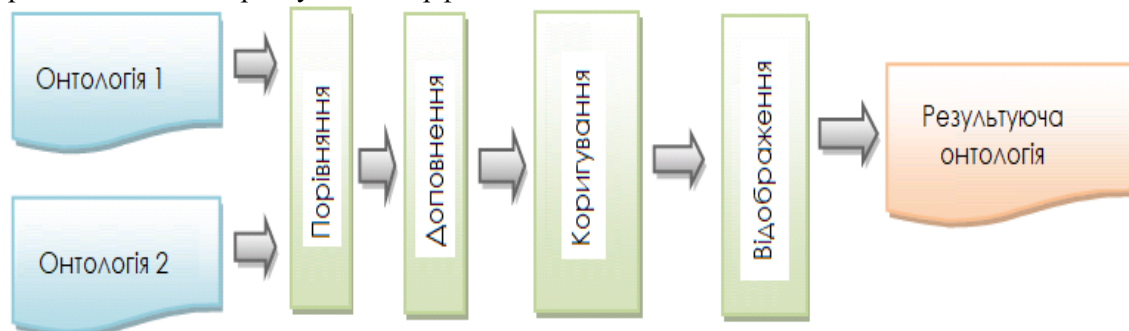


Рисунок 2 – Схема роботи інструменту порівняння і об'єднання онтологій

Запропоновано концептуальну модель програмної системи інтеграції даних, яка дозволяє користувачеві автоматично визначати і встановлювати семантичні зв'язки двох онтологій з подальшим встановленням взаємозв'язку схем інтегрованих систем, а також виконувати алгоритм формування запитів на додавання інформації з однієї системи в іншу.

Висновок

Розроблено математичну модель інтеграції даних інформаційних систем вузу з неоднорідними онтологічними специфікаціями, що дозволяє аналізувати семантичні зв'язки, що виникають між подібними елементами онтологій інтегрованих інформаційних систем.

Розроблена програмна система інтеграції даних інформаційних систем, що реалізує запропоновані моделі, методи і алгоритми.

Список використаних джерел

3. Mitra P., Wiederhold G.. Resolving the ontological heterogeneity in ontologies. In Proceedings of the ECAI'02 workshop on Ontologies and Semantic Interoperability, Lyon, 2002.
4. Mena E., Illarramendi A., Kashyap V., Sheth A.. OBSERVER: An approach for query processing in global information systems based on interoperation across preexisting ontologies. Distributed and Parallel Databases, An International Journal, N.2, 2015.
5. Guarino N. Formal ontology, conceptual analysis and knowledge presentation//International Journal of Human and Computer Studies, 43(5/6), pp. 625–640.