

УДК 004.89:37.041

Черемісіна Л. О.

НПУ ім. М. П. Драгоманова, Київ, Україна

ЗАСТОСУВАННЯ ОНТОЛОГІЙ ДЛЯ СТВОРЕННЯ БАЗ ЗНАНЬ НАВЧАЛЬНИХ КОМП'ЮТЕРНИХ СИСТЕМ

DOI: 10.14308/ite000515

У статті розглядається застосування онтологій для використання та розробки комп'ютерних навчальних систем. Розглянуто недоліки педагогічних програмних засобів та систем дистанційного навчання та переваги використання онтологій при їх розробці. Обґрунтовано актуальність створення навчальних комп'ютерних систем, заснованих на систематизованих знаннях. Розглянуто класифікацію властивостей, використання та переваг онтологій. Схарактеризовано підходи до вирішення проблеми відображення онтологій, перший з яких – ручне відображення, другий – зіставлення імен понять на основі їх лексичної подібності та використання спеціально розроблених словників. Зроблено аналіз мов, призначених для формального опису онтологій. Розглянута формальна математична модель онтологій та проблема узгодженості онтологій, яка полягає в тому, що різними розробниками для однієї й тієї ж предметної області можуть бути створені онтології, синтаксично або семантично гетерогенні, і для їх сумісного використання необхідна трансляція або відображення. Запропоновано алгоритм об'єднання онтологій. Схарактеризовано практичне значення розробки онтології для електронних освітніх ресурсів та перспективи подальших розробок і досліджень, а саме реалізація інших складових системи інтеграції, формалізація процесів інтеграції та розробка більш універсальних алгоритмів розширення онтологій програмного забезпечення.

Ключові слова: штучний інтелект, база знань, онтологія, навчальні системи.

Обсяги інформації з кожним роком стрімко збільшуються, відбувається переважання інформацією. Потік інформації надходить з науки, бізнесу, Інтернету та інших джерел. Але знання є не лише у людини, вони містяться також у накопичених даних, які необхідно аналізувати. Такі знання часто називають «прихованими».

Очевидно, що для виявлення «прихованих» знань потрібно застосовувати спеціальні методи автоматичного аналізу даних, за допомогою яких доводиться «добувати» знання із величезного об'єму інформації.

Електронні освітні ресурси (ЕОР) активно розробляються та впроваджуються в навчальний процес [1-2]. Сучасні програмні засоби (ПЗ), в тому числі і системи дистанційного навчання (СДН), успішно вирішують завдання створення, зберігання і «доставки» користувачеві знань, але при цьому існує ряд недоліків:

1. При розробці ПЗ та СДН не враховуються потреби та індивідуальні параметри того, кого навчають. Заданий план навчання поширюється на всіх студентів однаково, незалежно від навичок і знань конкретного індивіда. Строго задана траєкторія навчання, заздалегідь поданий певний обсяг освітньої інформації, не дозволяють врахувати особливості кожного споживача знань, що істотно відображається на якості освіти.
2. Підтримка та оновлення ПЗ та СДН – шлях дорогої і трудомісткий. Поява нових знань в предметній галузі неминує призведе до оновлення всіх освітніх ресурсів, в яких використовується це знання. Крім того, використання одних і тих же знань у різних предметних галузях робить вкрай нелегким підтримку і оновлення освітнього контенту до актуальної версії.

Сьогодні для вирішення перерахованих вище проблем найбільш доцільним є застосування баз знань, що представляють собою модель або концепцію зберігання знань. Повноцінні бази знань містять в собі не тільки фактичну інформацію, але й правила виведення, допускають автоматичні висновки про нововведені факти і як наслідок, – осмислену обробку інформації.

Ефективним засобом представлення та систематизації знань є онтології. Онтології використовуються для формальної специфікації понять і відношень, які характеризують певну галузь знань. Перевагою онтологій як способу представлення знань є їх формальна структура, яка спрощує комп'ютерну обробку інформації.

Онтологічний підхід набуває все більшого використання для найрізноманітніших типів інформаційних систем (ІС) та ЕОР [3-6]. Про зростаючу роль онтологій для ІС свідчить також введення в роботі [7] поняття керованих онтологіями інформаційних систем (Ontology-Driven Information Systems – ODIS). Онтології можуть використовуватися в ODIS під час їх розробки, використання, інтеграції, а також можуть підтримувати різні частини інформаційної системи – інтерфейси користувача, бази даних чи інші компоненти.

КНС, що використовують онтології, знаходять застосування при розробці та моделюванні систем представлення та інтерпретації інформації в процесі навчання. Онтологія визначає терміни, за допомогою яких можна описати і структурувати предметну область. Використання онтологій ефективно під час пошуку і об'єднання інформації з різних джерел і середовищ. Мова онтологій використовується для надання інформації чітко визначеного значення і являє собою загальний набір термінів для опису та подання предметної області, що вивчається.

Вона використовується як джерело даних для багатьох комп'ютерних додатків (для інформаційного пошуку, аналізу текстів, вилучення знання), дозволяючи більш ефективно обробляти складну і різноманітну інформацію. Вирішальний вплив на функціональні можливості освітнього контенту має модель даних, яка використовується для представлення знань. Перевагою онтологій в якості способу представлення знань є їх формальна структура, яка спрощує їх комп'ютерну обробку.

Таким чином, актуальною є задача створення комп'ютерних навчальних систем (КНС), заснованих на систематизованих знаннях.

Онтологія – представлення деякою мовою програмування знань про певну предметну область. Однак, не існує загальноприйнятого визначення «онтології», різняться погляди і на те, що вони мають включати в себе [8-9].

Щодо методів побудови онтологій, то на сьогодні існує ряд мов, призначених для формального опису онтологій. Серед найбільш відомих і використовуваних: KIF (Knowledge Interchange Format), DAML+OIL (DARPA Agent Mark up Language), OWL (Ontology Web Language). Існують також інструментальні засоби, що підтримують розробку онтологій відповідно до цих специфікацій.

Аналізуючи конструкції даних мов формального опису, можна помітити, що навіть у найбільш розширеній з них OWL, що включає OIL+DAML, існують детальні можливості лише для задавання класів, підкласів та їх членів (таксономії), для інших же типів відношень не передбачаються спеціальні елементи – їх можна задавати лише через властивості класів. На практиці більшість із вже створених онтологій є максимум ієрархічною структурою понять предметної області. Тобто розробниками розглядаються лише такі відношення між поняттями, як «вид-клас» та рідше «об'єкт-атрибут».

Тоді, як у деяких роботах [10], розроблено класифікацію властивостей онтологій та уточнено перелік структурних властивостей (таксономічні зв'язки, композиційні зв'язки, топологічні зв'язки, зв'язки сутностей з процесами, причинно-наслідкові зв'язки, часові та просторові зв'язки).

Відповідно до концепції Semantic Web створення RDF-описів та OWL-онтологій покладене на окремих розробників. І на сьогодні, попри досить незначні напрацювання в плані розробки онтологій, вже виникає проблема узгодженості онтологій, яка полягає в тому,

що різними розробниками для однієї й тієї ж предметної області можуть бути створені онтології, синтаксично або семантично гетерогенні, і для їх сумісного використання необхідна трансляція або відображення (виявлення відповідності між поняттями двох онтологій) [11].

Існує декілька підходів до вирішення проблеми відображення онтологій, перший з яких – ручне відображення, шляхом встановлення відношень між концептами, здійснювалося для деяких великих онтологій [12]. Проблема застосування ручного відображення в тому, що розмір онтологій може бути дуже великим і продовжуватиме нарощуватися, що вимагатиме надзвичайно багато людських зусиль для їх відображення. Тому, природно, що дослідники шукають шляхи відобразити онтології автоматично.

Досить значна кількість досліджень, серед яких і [13], присвячені розробці засобів відображення онтологій на основі методів машинного навчання, серед яких особливою популярністю користуються методи класифікації текстів. Однак результати тут залежать від якості навчальних даних, а підготовка їх вручну для сотень понять досить складна і дорога, що зменшує привабливість текстової класифікації.

Ще одним підходом до відображення онтологій є зіставлення імен понять на основі їх лексичної подібності та використання спеціально розроблених словників (Word Net), в яких описані відношення між концептами (синонімія) та властивості ряду концептів.

Альтернативним напрямком досліджень є автоматичне створення онтологій, яке на сьогодні зводиться до автоматичного анування текстів у Web. Аналіз робіт в цьому напрямку подано в [14], де показано їх обмеження виділенням певного типу відношень для анування, або ж використанням для анування певної онтології. Результати аналізу пропонується автоматично відобразити у OWL – описи за допомогою OntoSem2OWL.

Розглянемо математичну модель онтології.

Під формальною моделлю онтології O розумітимемо впорядковану трійку такого вигляду: $O = (C, R, F)$,

де C – скінченна множина концептів (понять, термінів) предметної області, яку задає онтологія O ;

R – скінченна множина відношень між концептами (поняттями, термінами) заданої предметної області;

F – скінченна множина функцій інтерпретації (аксіоматизація), заданих на концептах чи відношеннях онтології O .

Зазначимо, що природним обмеженням, що накладається на множину C , є його скінченність і не порожність. Інша річ з компонентами F і R у визначенні онтології O . Зрозуміло, що і в цьому випадку F і R мають бути скінченними множинами.

Завдання об'єднання онтологій є найбільш часто вирішуваним при створенні бази знань КНС. Тому розглянемо більш детально роботу підсистеми об'єднання онтологій. Онтологія розглядається як граф класів. Кожне поняття є вершиною, а відносини клас-підклас є дугами. Клас без базового класу є вершиною графу. Дозволено множинне спадкування. Будемо розглядати онтології, що представляються єдиним графом понять з однією кореневою вершиною. Граф не має посилань на зовнішні онтології.

Алгоритм об'єднання онтологій включає наступні укрупнені дії.

1. Вирівнювання порядків онтологій.
2. Розв'язання конфлікту імен.
3. Об'єднання класів та їх властивостей і відносин.

Під порядком онтології розуміється кількість рівнів ієрархії графу. Пропонований алгоритм може працювати при виконанні двох вимог, які пред'являються до онтологій, що об'єднуються:

- кількість рівнів ієрархії повинна бути однаковою;
- не повинно бути збігу імен понять, що знаходяться в онтологіях на різних рівнях ієрархії, тобто повинен бути відсутнім конфлікт імен.

В алгоритмі одна з онтологій вважається базовою, інша – тою, що доповнює. Алгоритм буде працювати більш ефективно, якщо базова онтологія більше, ніж та, що доповнює. Імовірність того, що онтології, які об'єднуються, будуть мати різну ієрархічну структуру, досить велика. Для коректної роботи запропонованого алгоритму необхідно вирівнювання порядків онтологій. Вирівнювання порядків – це процес приведення ієрархічної структури онтології з меншою кількістю рівнів ієрархії до структури другої онтології. Процес реалізується шляхом додавання в ієрархічну структуру фіктивних рівнів ієрархії. Імена понять на фіктивних рівнях ієрархії не розглядаються при роботі алгоритму.

Сутність запропонованого алгоритму полягає в наступному.

1. Формування списку збіжних імен понять онтологій, що знаходяться на різних рівнях ієрархії.
2. Формування списку збіжних імен понять онтологій, що знаходяться на однакових рівнях ієрархії.
3. Формування списку імен понять онтології, що доповнює, для яких відсутня синонімічність понять у базовій онтології на однакових рівнях ієрархії.

Список, отриманий на кроці 1 до початку об'єднання онтологій, має бути розглянутий експертом з метою розв'язання існуючого конфлікту імен. Тобто конфлікт імен повинен бути усунений до роботи алгоритму на кроці 2 і 3. На кроці 2 синонімічні поняття онтологій аналізуються модулем «перевірка відповідності». На кроці 3 незбіжні поняття обробляються модулем «аналіз елементів, що не збігаються».

Запропонований алгоритм об'єднання онтологій є важливою складовою частиною інструментальних засобів інтеграції ЕОР для створення онтологічних баз знань. Оскільки традиційний підхід до навчання в наш час суттєво потіснили альтернативні методи навчання, які базуються на інформаційних технологіях. Сучасні КНС є засобом представлення інформації, засвоєння знань і умінь, проміжної і підсумкової перевірки. Онтологічний підхід до створення предметного наповнення ЕОР може бути досить продуктивним, оскільки за його застосування до змісту навчання кількох предметів можна отримати абсолютно нове бачення міжпредметних зв'язків.

Ігнорування системних аспектів інженерії програмного забезпечення, як правило, веде до того, що ПЗ і СДН не відповідають вимогам обраного обладнання, або не інтегрується іншими системами. Застосування онтологій для представлення змісту навчальних комп'ютерних дозволить автоматизувати допомогу викладачам при створенні відповідних навчально-методичних матеріалів.

Завданням подальших розробок і досліджень має бути реалізація інших складових системи інтеграції, формалізація процесів інтеграції та розробка більш універсальних алгоритмів розширення онтологій програмного забезпечення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Лапінський В. В., Міна А. С., Скрипка К. І. Міжнародні тенденції розвитку інформатизації освіти та підвищення її якості // Інформаційні технології і засоби навчання – 2010. – № 5 (19). – Режим доступу: <http://lib.iitta.gov.ua/230/>.
2. Биков В. Ю., Лапінський В. В. Методологічні та методичні основи створення і використання електронних засобів навчального призначення // Комп'ютер в школі і сім'ї – 2010. – № 2 (98). – С. 3-6.
3. Вороной О.С., Єгошина Г.А. Засоби інтеграції онтологій предметних областей для створення баз знань інтелектуальних навчальних систем// Штучний інтелект. – 2010. – № 2. – С. 124-130.
4. Келеберда И.Н., Лесная Н.С., Репка В.Б. Использование мультиагентного онтологического подхода к созданию распределенных систем дистанционного обучения // Educational Technology & Society. – 2004. – № 7 (2).
5. Козак І. А. Централізований підхід до опису Web-онтологій // Штучний інтелект. – 2008. – № 4. – С. 80-85.
6. Палагин А. В., Величко В. Ю., Стрижак А.Е., Попова М.А. Инструменты поддержки процессов аналитической деятельности эксперта при тематическом исследовании

- информационных ресурсов и источников // International Journal “Information Technologies and Knowledge”, Vol. 4, Number 4, 2010. С. 329 – 347.
7. Guarino, N. Formal Ontology and Information Systems, in N. Guarino (Ed.) Formal Ontology in Information Systems. – Amsterdam, Netherlands: IOS Press (1998). – P. 3-15.
 8. Клещев А.С., Артемьева И.Л. Математические модели онтологий предметных областей. Часть 1. Существующие подходы к определению понятия «онтология» / Научно-техническая информация. – Серия 2 «Информационные системы и процессы». – 2001. – № 2. – С. 20-27.
 9. Пальчунов Д.Е. Моделирование мышления и формализация рефлексии I. Теоретико-модельная формализация онтологии и рефлексии // Философия науки. – 2006. – № 4 (31). – С. 86-114.
 10. Шалфеева Е.А. Классификация структурных свойств онтологий // Искусственный интеллект. – 2005. – № 3. – С. 67-77.
 11. Dou D., McDermott D. Qi. P. Ontology Translation on the Semantic Web // In Proceedings of International Conference on Ontologies, Databases and Applications of Semantics (ODBASE 2003). LNCS 2888. – Berlin Heidelberg. Springer-Verlag. – 2003. – P. 952-969.
 12. Niles I., Pease A. Mapping WordNet to the SUMO Ontology // Proc of the IEEE International Knowledge Engineering conference (2003).
 13. Zhongli Ding, Yun Peng, Rong Pan, Yang Yu. A Bayesian Methodology towards Automatic Ontology // Proceedings of the AAAI-05 C&O Workshop on Contexts and Ontologies: Theory, Practice and Applications. -July 09, 2005.
 14. Akshay Java, Sergei Nirneburg, Marjorie McShane, Timothy Finin, Jesse English, Anupam Joshi. Using a Natural Language Understanding System to Generate Semantic Web Content // Final version to appear, International Journal on Semantic Web and Information Systems. – 2007. – 3 (4). – Режим доступа: <http://www.igi-pub.com/>.

Стаття надійшла до редакції 10.11.14

Cheremisina Lyubov

NPU. M. P. Drahomanova, Kiev, Ukraina

USE OF ONTOLOGIES FOR KNOWLEDGE BASES CREATION TUTORING COMPUTER SYSTEMS

This paper deals with the use of ontology for the use and development of intelligent tutoring systems. We consider the shortcomings of educational software and distance learning systems and the advantages of using ontology's in their design. Actuality creates educational computer systems based on systematic knowledge. We consider classification of properties, use and benefits of ontology's. Characterized approaches to the problem of ontology mapping, the first of which – manual mapping, the second – a comparison of the names of concepts based on their lexical similarity and using special dictionaries. The analysis of languages available for the formal description of ontology. Considered a formal mathematical model of ontology's and ontology consistency problem, which is that different developers for the same domain ontology can be created, syntactically or semantically heterogeneous, and their use requires a compatible broadcast or display. An algorithm combining ontology's. The characteristic of the practical value of developing an ontology for electronic educational resources and recommendations for further research and development, such as implementation of other components of the system integration, formalization of the processes of integration and development of a universal expansion algorithms ontology's software.

Key words: artificial intelligence, knowledge base, ontology, tutoring system.

Черемисина Любовь Александровна

НПУ им. М. П. Драгоманова, Киев, Украина

ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЙ ДЛЯ СОЗДАНИЯ БАЗ ЗНАНИЙ УЧЕБНЫХ КОМПЬЮТЕРНЫХ СИСТЕМ

В статье рассматривается применение онтологий для использования и разработки интеллектуальных обучающих систем. Рассмотрены недостатки учебных педагогических программных средств и систем дистанционного обучения, а также преимущества

использования онтологий при их разработке. Обоснована актуальность создания интеллектуальных обучающих систем основанных на систематизированных знаниях. Рассмотрено преимущества использования и классификацию свойств онтологий. Охарактеризованы подходы к решению проблемы отражения онтологий, первый из которых – ручное отображения, второй – сопоставление имен понятий на основе их лексического сходства и использования разработанных словарей. Сделан анализ языков, предназначенных для формального описания онтологий. Рассмотрена формальная математическая модель онтологий и проблема согласованности онтологий, которая заключается в том, что различными разработчиками для одной и той же предметной области могут быть созданы онтологии, синтаксически или семантически гетерогенные, и для их совместного использования необходима трансляция или отображения. Предложен алгоритм объединения онтологий. Охарактеризовано практическое значение разработки онтологии для электронных образовательных ресурсов и перспективы их дальнейших разработок и исследований, а именно реализация составляющих системы интеграции, формализация процессов интеграции и разработка более универсальных алгоритмов расширения онтологий программного обеспечения.

Ключевые слова: искусственный интеллект, база знаний, онтология, учебные системы.