

УДК 004.896

С.В. Титенко

ПОБУДОВА ДИДАКТИЧНОЇ ОНТОЛОГІЇ НА ОСНОВІ АНАЛІЗУ ЕЛЕМЕНТІВ ПОНЯТІЙНО-ТЕЗИСНОЇ МОДЕЛІ

Вступ

Задача побудови семантичної моделі предметної області навчання відіграє ключову роль у процесі створення інтелектуальної навчальної системи. Таку модель прийнято називати онтологією. Відношення дидактичного слідування, що вказує на те, що певне поняття дидактично передує іншому, і є ключовим семантичним зв'язком для понять навчальної системи [1–6]. На основі такого відношення можна зобразити дидактичну структуру навчальних понять у вигляді спеціальної семантичної мережі, яку називатимемо дидактичною онтологією (ДО). Наявність ДО дає можливість застосовувати різноманітні технології індивідуалізації навчання: адаптивну навігацію, адаптацію контенту [7] та ін. Натомість, формування подібних відношень за допомогою ручної роботи експерта [1] в системах з відкритим типом контенту [8, 9] ускладнюється через значну трудомісткість і необхідність підтримки великої кількості об'єктованої навчальної інформації. Під системами з відкритим типом контенту прийнято розуміти такі системи, контент яких невідомий на етапі проектування і програмної реалізації системи та передбачається, що він буде додаватися і розширюватися на етапі використання [9]. Сюди слід віднести як розподілені WWW-ресурси, так і навчальні системи з розвиненими можливостями щодо адміністрування контенту і створення великих сховищ навчальних матеріалів [8]. Саме такі системи можуть забезпечити сучасні потреби електронного навчання.

У праці [3] пропонується метод автоматичного визначення дидактичної черговості понять у прикладах програмних кодів з мовою програмування С. Цей метод ґрунтується на послідовності лекцій і їх зв'язку з програмними прикладами. Як зазначають автори, запропонований принцип може бути застосований для інших предметних областей завдяки технологіям інформаційного пошуку. Хоч ця ідея вбачається перспективною, однак конкретні приклади й досі не були розглянуті. Крім того, інфор-

мація про черговість лекцій не завжди доступна у випадку систем, метою яких є генерація навчального середовища. До таких систем належать системи підтримки безперервного навчання [7]. Таким чином, проблема автоматичної побудови дидактичної онтології є актуальною задачею для галузі сучасних систем навчання. У попередніх працях [10, 11] було запропоновано понятійно-тезисну модель формалізації дидактичного тексту (ПТМ). ПТМ застосовувалася як засіб для генерації тестових завдань [11], а також як засіб для побудови термінологічного довідника дистанційного курсу [12]. В даному дослідженні ставилося завдання використати ПТМ для автоматичної побудови дидактичної онтології.

Постановка задачі

Метою статті є розробка методу автоматичної побудови дидактичної онтології в навчальній системі на основі даних понятійно-тезисної моделі. Формально під дидактичною онтологією будемо розуміти орієнтований ациклічний граф, вершинами якого виступають навчальні поняття, а дуги вказують на відношення дидактичного слідування між ними. Як вершини графа використовуються поняття, описані в понятійно-тезисній базі навчального курсу, а дуги мають бути визначені на основі семантико-синтаксичного аналізу елементів ПТМ. Для аналізу елементів ПТМ і виявлення відношень дидактичного слідування між поняттями пропонується застосувати математичний апарат нечітких правил [13].

Понятійно-тезисна модель як підґрунтя для побудови дидактичної онтології

Структуру, сутності і формальний опис ПТМ наведено в публікаціях [10, 11]. Елементи ПТМ є результатом формалізації дидактичного тексту, виконаної експертом. ПТМ служить для розв'язання цілого ряду задач у навчальній системі, серед яких автоматизація контролю знань, автоматизація побудови термінологічного довідника курсу тощо. Тому розширення області застосування ПТМ для вирішення проблеми побудови ДО ефективно з точки зору оптимального використання трудових ресурсів. До ПТ-елементів належать поняття і тези, які є дидактичним вираженням знань з предметної області у формі фрагментів навчального тексту. Синтаксичний аналіз цього тексту із врахуван-

ням семантики, закладеної в класифікації ПТ-елементів, дає можливість автоматично визначити дидактичні співвідношення між поняттями. Ідея полягає у виявленні понять, які синтаксично входять в інші поняття або в їх тези. На основі таких даних можна зробити висновки про дидактичне слідування. У табл. 1 наведемо деякі приклади ПТ-елементів, які ілюструють подібну ситуацію (інформація з навчального курсу “Використання технології ADO в Delphi” з циклу дисциплін напрямку “Програмна інженерія”).

Таблиця 1. Фрагмент понятійно-тезисної бази по курсу “Використання технології ADO в Delphi”

Поняття	Теза
ADO	Інтерфейс високого рівня для роботи з OLE DB, орієнтований на використання в прикладних програмах
	Більш широко інтерпретує поняття “дані” порівняно з BDE
	Постачається в складі MDAC
OLE DB	Являє собою інтерфейс системного рівня і призначена для використання, в першу чергу, системними програмістами
Microsoft	У середині 90-х років приступила до заміни технології ODBC технологією OLE DB
MDAC	Компоненти доступу до даних Microsoft
	Є технологією Microsoft доступу до баз даних і містить у собі ADO, OLE DB, ODBC і RDS
Базовий об’єкт ADO Field	Зберігає всю необхідну інформацію про одне поле набору даних

Таким чином, у випадку, коли в тезі поняття “1” зустрічається поняття “2”, то з деякою мірою впевненості можна стверджувати, що для успішного засвоєння поняття “1” потрібно спочатку засвоїти поняття “2”. Ще з більшою мірою впевненості можна стверджувати, що поняття “2” дидактично передують поняттю “1”, якщо поняття “1” містить поняття “2” у своїй назві.

Натомість, можливі випадки, коли теза поняття “1”, в якій знайдено поняття “2”, має такий характер, що з дидактичної точки зору буде лише підкреслювати той факт, що з по-

няття “1” впливає поняття “2”, а не навпаки. Такі випадки характерні тоді, коли в деякому навчальному фрагменті тексту ключовим предметом обговорення виступає саме поняття “2”, а поняття “1” з’являється в контексті як другорядне. Тут автором навчального тексту передбачається, що поняття “1” має бути відоме учневі, і тому воно може виступати в даному тексті як фонове.

Аналіз елементів понятійно-тезисної моделі і побудова дидактичної онтології на основі нечітких правил

Враховуючи сказане вище, а також той факт, що одного поняття стосується в загальному випадку більше однієї тези і цілком можлива ситуація, коли поняття проникають у тези одне одного, зазначимо, що задача автоматичного визначення дидактичної послідовності понять на основі ПТМ має неоднозначний характер. У зв’язку з цим для розв’язання задачі пропонуємо застосувати апарат нечітких правил на основі фактора впевненості [13].

Базовими правилами для визначення дидактичного слідування понять будуть такі правила.

Правило № 1. Якщо поняття “1” фігурує в назві поняття “2”, то поняття “1” дидактично передують поняттю “2” з високим ступенем достовірності.

Правило № 2. Якщо поняття “1” фігурує в тезі поняття “2”, то поняття “1” дидактично передують поняттю “2” з деякою достовірністю.

Правило № 3. Також для деяких випадків діятиме зворотне правило: якщо поняття “1” фігурує в тезі поняття “2”, то поняття “2” дидактично передують поняттю “1” з деякою достовірністю.

Опишемо множини і відношення, які одержуються в результаті синтаксичного аналізу тез T і понять C . Зв’язок між тезами і поняттями, які були знайдені в цих тезах завдяки синтаксичному аналізу, задається відношенням

$$CinT : T \rightarrow 2^C.$$

Зв’язок між базовими і похідними поняттями передбачає, що назва базового поняття фігурує в назві похідного поняття. Цей зв’язок задається відношенням

$$CinC : C \rightarrow 2^C.$$

Фактор впевненості для класу тези вказує на міру достовірності того, що поняття, знайдене в тезі такого класу, буде дидактично пе-

редумовою для поняття, якому дана теза належить. Множина класів тез відображається на значенні фактора впевненості:

$$TClassCF : TClass \rightarrow X,$$

де $X = [-1...1]$.

При цьому, якщо $TClassCF(tClass_i) < 0$, то вважатимемо, що тези даного класу свідчать про те, що знайдені в них поняття не передують, а дидактично слідує за поняттям, якому належить теза, з фактором впевненості $|TClassCF(tClass_i)|$ (див. вище базове правило № 3). Такі класи тез називатимемо *реверсними*.

Щоб далі описувати правила, будемо застосовувати предикат, що служить для опису твердження про те, що деяке поняття c_k дидактично передує деякому поняття c_l :

$$concept_before(c_k, c_l).$$

Отже, базове правило № 1, виражене в символній формі, матиме такий вигляд:

$$c_k \in CinC(c_l) \rightarrow concept_before(c_k, c_l) \langle CFeinc \rangle,$$

де фактор впевненості $CFeinc$ дорівнює 0,99.

Базове правило № 2, виражене в символній формі, матиме вигляд

$$t \in TC(c_l) \wedge c_k \in CinT(t) \wedge \\ \wedge TClassCF(TClass(t)) > 0 \rightarrow$$

$$\rightarrow concept_before(c_k, c_l) \langle TClassCF(TClass(t)) \rangle.$$

Тут $TClassCF(TClass(t))$ дає значення фактора впевненості залежно від класу тези. Базове правило № 3, виражене в символній формі, матиме вигляд

$$t \in TC(c_l) \wedge c_k \in CinT(t) \wedge \\ \wedge TClassCF(TClass(t)) < 0 \rightarrow$$

$$\rightarrow concept_before(c_l, c_k) \langle -TClassCF(TClass(t)) \rangle.$$

У табл. 2 подається набір базових класів тез ПТМ і відповідних їм факторів впевненості, які вибрані евристично із врахуванням семантико-дидактичних особливостей тез відповідних класів.

Після того як здійснено синтаксичний аналіз ПТ-елементів і побудовано множини $CinC$ і $CinT$, тобто побудовано базу фактів, необхідно підрахувати сукупний фактор впевне-

ності для гіпотез про дидактичні відношення між парами понять. Слід підрахувати сукупний фактор впевненості для гіпотез на основі правил, які приводять до однакового результату із врахуванням і усуненням можливих протиріч.

Визначимо множину понять, які, ймовірно, є такими, що дидактично передують даному поняттю. Цю множину понять називатимемо *вхідними поняттями-кандидатами* даного поняття $a \in C$:

$$TryCtoC(a) = \{c \in C : (c \in CinC(a)) \vee$$

$$\vee (c \in CinT(t) \wedge t \in TC(a) \wedge TClassCF(TClass(t)) > 0) \vee$$

$$\vee (t \in TC(c) \wedge a \in CinT(t) \wedge TClassCF(TClass(t)) < 0)\}.$$

Після отримання вхідних понять-кандидатів наступним кроком буде підрахунок фактора впевненості гіпотези про те, що окреме вхідне поняття-кандидат $c \in TryCtoC(a)$ дидактично передує даному поняттю a . Спочатку знаходимо сукупність факторів впевненості тих правил, результатом яких буде твердження $before_concept(c, a)$, яке вказує на дидактичну послідовність понять c і a :

$$CFs(c, a) = \{x : (x = CFeinc \wedge c \in CinC(a)) \vee$$

$$\vee (x = TClassCF(t) \wedge t \in TC(a) \wedge c \in$$

$$\in CinT(t) \wedge TClassCF(TClass(t)) > 0) \vee$$

$$\vee (x = -TClassCF(t) \wedge t \in TC(c) \wedge a \in$$

$$\in CinT(t) \wedge TClassCF(TClass(t)) < 0)\}.$$

Далі, щоб отримати сукупний фактор впевненості для $before_concept(c, a)$, необхідно об'єднати одержані значення $CFs(c, a)$, для чого застосуємо послідовне попарне об'єднання за принципом, поданим у [13]:

$$CF = CF + CF_i - CF \cdot CF_i,$$

де $CF_i \in CFs(c, a)$, $i = 2, \dots, n$.

Таким чином, отримуємо фактор впевненості для твердження $before_concept(c, a)$, що вказує на те, що поняття c дидактично передує поняттю a з фактором впевненості $CF(before_concept(c, a))$. На основі цих даних будується матриця дидактичних зв'язків між поняттями $CF_CtoC = \|cf_CtoC_{ij}\|$, що задає відповідне відношення $CtoC$:

$$CtoC : C \times C.$$

Стовпці і рядки матриці CF_CtoC відповідають поняттям, а значення елементів – значенню фактора впевненості того, що поняття c_j (рядок) дидактично передує поняттю c_i (стовп-

чик) і становить $CF(\text{before_concept}(c_i, c_j))$. Для позначення збереженого в матриці CF_CtoC фактора впевненості $CF(\text{before_concept}(c_i, c_j))$ застосовуватимемо оператор $CF_CtoC(c_i, c_j)$, результатом якого буде значення cf_CtoC_{ij} .

Таблиця 2. Базові класи тез ПТМ і відповідні їм фактори впевненості

Клас тези $tClass \in$ $\in TClasses$	Пояснення	Фактор впевненості $TClassCF(tClass)$
$tDefinition$	Визначення	0,95
$tDestination$	Призначення	0,8
$tEssence$	Сутність	0,65
$tSyntax$	Синтаксис	0,8
$tGeneral$	Загальна відомість	0,2
$tList$	Список	0,5
$tListItem$	Елемент списку	0,5
$tImage$	Зображення	0
$tSynonym$	Синонім	0,95
$tAbbrev$	Абревіатура	0
$tAbbrevDecode$	Розшифровка абревіатури	0,95
$tCode$	Програмний код	0,6
$tAttaching$	Прикріплення	0
$tReverseEssence$	Сутність вкладеного поняття	-0,5
$tReverseGeneral$	Загальна відомість про вкладене поняття	-0,2

Після підрахунку $CF(\text{before_concept}(c, a))$ для кожного із вхідних понять-кандидатів $c \in TryCtoC(a)$ необхідно знайти CF відношень між поняттям a і його вихідними поняттями-кандидатами:

$$TryCfromC(a) = \{c \in C : (a \in C \text{ in } C(c)) \vee$$

$$\vee (a \in C \text{ in } T(t) \wedge t \in TC(c) \wedge TClassCF(TClass(t)) > 0) \vee$$

$$\vee (t \in TC(a) \wedge c \in C \text{ in } T(t) \wedge TClassCF(TClass(t)) < 0)\}.$$

Аналогічно знаходиться $CF(\text{before_concept}(a, c))$ для всіх $c \in TryCfromC(a)$. Відповідним чином доповнюється матриця CF_CtoC . Під час цих обчислень з'являється можливість виявлення протиріч, коли відбуватиметься одночасне виконання нерівностей $CF_CtoC(a, c) > 0$ і $CF_CtoC(c, a) > 0$. В такому випадку, адаптуючи

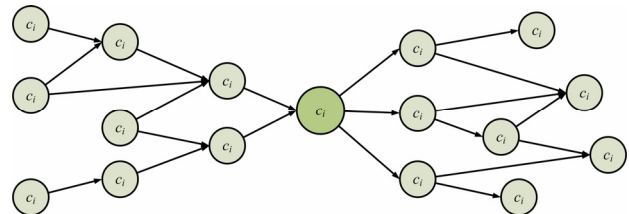
формулу з [13], обчислюємо нове значення фактора впевненості:

$$CF = \frac{\max(CF_CtoC(a, c), CF_CtoC(c, a))}{1 - \min(CF_CtoC(a, c), CF_CtoC(c, a))} - \frac{\min(CF_CtoC(a, c), CF_CtoC(c, a))}{1 - \min(CF_CtoC(a, c), CF_CtoC(c, a))}.$$

При цьому менше із значень $CF_CtoC(a, c)$, $CF_CtoC(c, a)$ прирівнюється нулю, тоді як замість більшого з них підставляється нове значення CF .

Візуалізація дидактичної онтології у вигляді дидактико-семантичних карт

Дидактичні зв'язки, визначені в процесі семантико-синтаксичного аналізу ПТ-елементів, дають можливість візуалізувати відношення між поняттями за допомогою побудови дидактико-семантичних карт понять. *Дидактико-семантична карта поняття* – це орієнтований ациклічний граф, вершинами якого виступають поняття, а кожна з дуг вказує на те, що поняття-початок дуги дидактично передують поняттю-кінцю дуги. При цьому так звані центром графа є поняття, для якого будується карта (рисунок). Карта послідовно показує всі зв'язки, починаючи від центрального поняття з дидактично передуючими поняттями (ліва частина графа) і всі зв'язки, починаючи від центрального поняття, з дидактично наступними поняттями (права частина графа). Крім того, такий граф є зваженим, так що як вага дуги виступає значення відповідного фактора впевненості $CF_CtoC(c_1, c_2)$, де c_1, c_2 – відповідно початок і кінець дуги.



Приклад графа дидактико-семантичної карти поняття

Якщо під час побудови карти виявлено цикл, тобто граф – циклічний, то це вказує на суперечливість ПТ-бази, що є сигналом для експерта про необхідність корегування ПТ-елемен-

тів. Подібна ситуація візуалізується спеціальним чином, так що експерту вказується шлях, на якому виникає цикл. Дані поняття цього шляху слід проаналізувати для внесення необхідних правок. Редагування, як правило, передбачатиме зміну класів для деяких із тез на реверсні класи. Після внесення правок карта поняття має перевірятись повторно.

Висновки

Ключовою перевагою запропонованого методу побудови ДО є зменшення трудових витрат на встановлення дидактичних відношень між поняттями за рахунок автоматизації цього процесу. Інша перевага полягає у придатності методу до застосування в різних предметних областях, що було перевірено на навчальних матеріалах у галузях технічних, математичних, економічних та педагогічних наук [14]. До

переваг слід також віднести незалежність від даних про черговість лекцій, що дає можливість застосовувати цю технологію для дидактичної формалізації словників та довідників, матеріали яких впорядковуються не дидактично, а за алфавітом. Ця перевага є актуальною також у випадку формалізації набору статей з певної тематики, а це, в свою чергу, має велике значення для підтримки навчання в нових галузях, якісні навчальні курси якими досі перебувають на стадії формування. Крім того, ПТМ розрахована на роботу з контентом відкритого типу, а тому разом із моделлю Tree-Net [15] забезпечує розв'язок проблеми семантичної формалізації контенту, який додається до системи на етапі її експлуатації. Подальші дослідження зосереджені на побудові методів використання ДО для генерації індивідуального навчального середовища. Результати публікуються на сайті [16].

С.В. Титенко

ПОСТРОЕНИЕ ДИДАКТИЧЕСКОЙ ОНТОЛОГИИ НА ОСНОВЕ АНАЛИЗА ЭЛЕМЕНТОВ ПОНЯТИЙНО-ТЕЗИСНОЙ МОДЕЛИ

Рассмотрен метод построения дидактической онтологии (ДО) в обучающей системе на основе семантико-синтаксического анализа элементов понятийно-тезисной модели с применением нечеткой логики. Описаны базовые правила для определения отношения дидактического следования между понятиями и представлено их математическое описание. Предложен метод визуализации ДО на основе дидактико-семантических карт.

S.V. Tytenko

CONSTRUCTION OF DIDACTIC ONTOLOGY BASED ON THE ANALYSIS OF CONCEPT-THESIS MODEL ELEMENTS

The present paper considers the method of didactic ontology (DO) construction in the educational system. It is based on the semantic and syntactic analysis of concept-thesis model elements by using fuzzy logic. We represent some basic rules for determining the didactic sequence relation and provide its mathematical description. Crucially, we introduce the method of DO visualization on the basis of didactic-semantic maps.

1. *Brusilovsky P.* Developing adaptive educational hypermedia systems: From design models to authoring tools // T. Murray, S. Blessing and S. Ainsworth (eds.): *Authoring Tools for Advanced Technology Learning Environment*. — Dordrecht: Kluwer Academic Publishers, 2003. — P. 377–409.
2. *Brusilovsky P., Pesin L.* Adaptive navigation support in educational hypermedia: An evaluation of the ISIS-Tutor // *J. of Computing and Information Technology*. — 1998. — 6, N 1. — P. 27–38.
3. *Brusilovsky P., Yudelson M., Sosnovsky S.* An adaptive E-learning service for accessing interactive examples // J. Nall and R. Robson (eds.). *Proceedings of World Conference on E-Learning, E-Learn 2004*, Washington, DC, USA, November 1–5, 2004. — AACE, 2004. — P. 2556–2561.
4. *Weber G., Brusilovsky P.* ELM-ART: An adaptive versatile system for Web-based instruction // *International Journal of Artificial Intelligence in Education* 12 (4), Special Issue on Adaptive and Intelligent Web-based Educational Systems, 2001. — P. 351–384.
5. *Brusilovsky P., Schwarz E., Weber G.* A tool for developing hypermedia-based ITS on WWW // C. Frasson, G. Gauthier and A. Lesgold (eds.). *Intelligent Tutoring Systems: Proceedings of the Third International Conference, ITS'96*. — Berlin: Springer, 1996. — P. 261–269.

6. *Weber G., Specht M.* User modeling and Adaptive Navigation Support in WWW-Based Tutoring Systems. Workshop on Adaptive Systems and User Modeling on the WWW on UM97. – Chia Laguna, Sardinia: Springer. – 1997.
7. *Гагарін О.О., Титенко С.В.* Дослідження і аналіз методів та моделей інтелектуальних систем безперервного навчання // Наукові вісті НТУУ “КПІ”. – 2007. – № 6. – С. 37–48.
8. *Brusilovsky P. and Henze N.* Open corpus adaptive educational hypermedia // P. Brusilovsky, A. Kobsa and W. Neidl (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science. – Berlin, Heidelberg; New York: Springer-Verlag, 2007. – 4321. – С. 671–696.
9. *Brusilovsky P.* Adaptive Navigation Support for Open Corpus Hypermedia Systems (Invited talk) // W. Nejdl, J. Kay, P. Pu and E. Herder (eds.). Proceedings of 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH’2008), Hannover, Germany, July 29–August 1, 2008. – Germany: Springer-Verlag, 2008 – P. 6–8.
10. *Титенко С.В., Гагарін О.О.* Семантична модель знань для цілей організації контролю знань у навчальній системі // Сб. тр. междунар. конф. “Інтелектуальний аналіз інформації–2006”. – К.: Просвіта, 2006. – С. 298–307.
11. *Титенко С.В.* Генерація тестових завдань у системі дистанційного навчання на основі моделі формалізації дидактичного тексту // Наукові вісті НТУУ “КПІ”. – 2009. – № 1. – С. 47–57.
12. *Титенко С.В.* Генерація індивідуального навчального середовища на основі моделі професійних компетенцій у Web-системі безперервного навчання // Вісн. Східноукр. нац. ун-ту ім. Володимира Даля. – 2009. – № 1 (131). – С. 267–273.
13. *Buchanan B.G., Shortliffe E.H. et al.* Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley, 1984.
14. *Портал знань* – Web-портал, ПЗ якого реалізує запропонований метод побудови дидактичної онтології [Електронний ресурс]. – Режим доступу: <http://www.znannya.org>.
15. *Титенко С.В., Гагарін О.О.* Моделювання області знань в системі безперервного навчання на основі інтеграції моделі контенту Tree-Net і понятійно-тезисної моделі // VIII междунар. конф. “Інтелектуальний аналіз інформації ІАІ-2008”, Київ, 14–17 мая 2008 г.: Сб. тр. / Ред. кол.: С.В. Сирота (гл. ред.) и др. – К.: Просвіта, 2008. – С. 475–484.
16. *Лабораторія СЕТ* – лабораторія семантичних технологій в дистанційному навчанні [Електронний ресурс]. – Режим доступу: <http://www.setlab.net>.

Рекомендована Радою
теплоенергетичного факультету
НТУУ “КПІ”

Надійшла до редакції
28 жовтня 2009 року