

## FORMAL APPARATUS AND SOFTWARE FOR AUTOMATION OF WEB BASED INDIVIDUAL LEARNING ENVIRONMENT PRODUCTION

Tytenko Sergiy, Gagarin Alexander

National Technical University of Ukraine "Kyiv Polytechnic Institute"

*The paper presents formal description of models' set, methods and algorithms for automation of individual learning environment production in the interdisciplinary information and learning web portal.*

### ФОРМАЛЬНИЙ АПАРАТ ТА ПРОГРАМНІ ЗАСОБИ АВТОМАТИЗАЦІЇ ПОБУДОВИ ІНДИВІДУАЛЬНОГО НАВЧАЛЬНОГО WEB-СЕРЕДОВИЩА

Титенко С.В., Гагарін О.О.

Національний технічний університет України «Київський  
політехнічний інститут»

*Представлено формальний опис комплексу моделей, методів і алгоритмів для автоматизації побудови індивідуального навчального середовища в міждисциплінарному інформаційно-навчальному Web-порталі*

Бурхливий розвиток інформаційних ресурсів в Україні і в світі та поширення засобів доступу до них породжують гостру потребу в спеціалізованих програмних системах, що спростять пошук освітньої інформації, релевантної поточним навчально-професійним потребам користувачів. Реалізація такого програмного забезпечення вимагає розробки спеціалізованих формальних моделей опису і передачі освітнього контенту інформаційно-навчальних середовищ. Подібна задача вирішувалась в роботах Брусилівського П. та його учнів, Башмакова А.І та Башмакова І.А., Манако А.Ф., Мазурок Т.Л., Войченко О.П., Данилової О.В. та ін. Актуально залишається задача розробки формальних моделей і методів та їх реалізація в програмному забезпеченні систем з індивідуалізованим доступом користувачів до затребуваної професійно-навчальної інформації в Web-середовищі.

Для забезпечення моделювання контенту на предметному рівні в даній роботі пропонується **понятійно-тезисна модель** (ПТМ) [1]. Вона застосовується як засіб моделювання смислу контенту

інформаційно-навчального середовища, при цьому формалізація відбувається всередині фрагменту навчального тексту.

Наріжним каменем структури моделі є *поняття* – предмет обговорення, деякий об'єкт з предметної області, про який в навчальному матеріалі є знання. *Теза* – це деяка відомість або твердження про поняття. Якщо поняття вказують на предмет контенту, то тези є описово-смысловим наповненням бази знань, яке розкриває характер і властивості наявних понять. З кожним поняттям пов'язується множина тез. Формально теза є одним або декількома реченнями, у яких мова йде безпосередньо про відповідне поняття, проте саме поняття там синтаксично не фігурує. Наведемо приклади: теза про поняття «процедура» – «дозволяє розбити програму на підпрограми»; теза про поняття «клас» – «може мати в своїй структурі не тільки поля-властивості, а й методи, тобто функції і процедури». Множина понять:  $C = \{c_1, \dots, c_{n1}\}$ . Множина тез:  $T = \{t_1, \dots, t_{n2}\}$ . Зв'язок між тезами і поняттями:  $CT: T \rightarrow C$ ,  $TC: C \rightarrow 2^T$ .

Елементи ПТМ виділяються експертом безпосередньо із тексту навчального фрагменту за допомогою спеціалізованих засобів користувацького інтерфейсу. У результаті кожен фрагмент  $v_i$  може стати джерелом довільної кількості тез  $t_i$ , що задається відображенням:  $TV: V \rightarrow 2^T$ ,  $VT: T \rightarrow V$ . Поняття, які стосуються даної навчальної ділянки, та відповідно навчальний матеріал, якого стосується дане поняття, визначаються операторами:

$$CV(v) = \{c: TV(v) \cap TC(c) \neq \emptyset\}, \quad VC(c) = \{v: TV(v) \cap TC(c) \neq \emptyset\}.$$

*Класифікація тез і понять* служить для збереження в БЗ інформації про змістовий або лексичний характер того чи іншого поняття чи тези:  $TClass = T \rightarrow TClasses$ ,  $CClass = C \rightarrow CClasses$ .

На основі семантико-синтаксичного аналізу елементів ПТМ і стенфордської моделі нечіткого виведення [2] розроблено метод автоматичної побудови *онтології предметної області*, що ґрунтується на відношенні дидактичного слідування [1]. Відношення дидактичного слідування між двома поняттями вказує на те, що певне поняття дидактично передує іншому, тобто в структурі навчального матеріалу повинно подаватися раніше. Сам метод ґрунтується на трьох базових нечітких логічних правилах, у

відповідності до кожного з яких ставиться фактор впевненості  $CF$  [2]:

*Правило №1.* Якщо поняття «1» фігурує в назві поняття «2», то поняття «1» є дидактичною передумовою поняття «2» з високим ступенем достовірності:

$$c_k \in \text{Cin}C(c_l) \rightarrow \text{concept\_before}(c_k, c_l) \langle \text{CFinc} \rangle$$

*Правило №2.* Якщо поняття «1» фігурує в тезі поняття «2», то поняття «1» є дидактичною передумовою поняття «2» з деякою достовірністю:

$$t \in TC(c_l) \wedge c_k \in \text{Cin}T(t) \wedge T\text{Class}CF(T\text{Class}(t)) > 0 \rightarrow \\ \rightarrow \text{concept\_before}(c_k, c_l) \langle T\text{Class}CF(T\text{Class}(t)) \rangle$$

*Правило №3.* Також для деяких випадків діятиме зворотнє правило: якщо поняття «1» фігурує в тезі поняття «2», то поняття «2» є дидактичною передумовою поняття «1» з деякою достовірністю:

$$t \in TC(c_l) \wedge c_k \in \text{Cin}T(t) \wedge T\text{Class}CF(T\text{Class}(t)) < 0 \rightarrow \\ \rightarrow \text{concept\_before}(c_l, c_k) \langle -T\text{Class}CF(T\text{Class}(t)) \rangle$$

У випадку, коли для протилежних гіпотез одночасно має місце  $CF > 0$ , істинною приймається гіпотеза із більшим значенням  $CF$ , при цьому фактор впевненості перераховується за формулою:

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

**Структурна модель контенту та модель професійних компетенцій** створені для забезпечення опису інформаційно-навчальних ресурсів Web-порталу у їх співвідношенні з професійними завданнями та компетенціями [3]. На цій основі розроблена **підсистема організації індивідуалізованого навчання**, яка служить для побудови індивідуального навчального середовища (ІНС). У залежності від цілей користувача і типу його освітнього запиту навчальний процес може приймати різні за цільовим призначенням і обсягом форми: 1) здобуття спеціальності:  $L\text{Exp}Aims(l_i) = Eq\text{Exp}$ ; 2) здобуття компетенції або адаптованої спеціальності:  $LSAims(l_i) = EqS$ ; 3) вивчення індивідуального навчального курсу:  $LVAims(l_i) = EqV$ ; 4) дослідження предметної області:  $LGAims(l_i) = EqG$ ; 5) вивчення окремого навчального поняття:  $LCAims(l_i) = EqC$ .

*Генерація ІНС* на здобуття спеціальності відбувається на основі відомостей про профіль спеціаліста. Повний набір компетенцій, що стосуються даного профілю, є *декомпозицією профілю спеціаліста*, який визначатимемо наступним чином:

$$SDExp(exp) = \{s \in S: s \in SExp(exp) \vee s \in DescS(a), \text{ де } a \in SExp(exp)\}.$$

Усю сукупність контенту декомпонованого профілю спеціаліста називатимемо *профільною областю контенту* даного спеціаліста:  $VSDExp(exp) = \{v: v \in VatS(s), \text{ де } s \in SDExp(exp)\}.$

Після отримання сукупності контенту  $V' = VSDExp(exp)$ ,  $V' \sqsubseteq V$  декомпозиції профілю  $exp \in Exp$  для його ієрархічного структурування застосовуються базові відношення ієрархічності між елементами контенту  $F$  і  $Ch$ . У результаті отримуємо деяку сукупність піддерев контенту, які можуть розглядатися у якості *набору індивідуальних навчальних курсів і модулів*. Оператор  $Roots(V')$  вкаже на корені новоутворених піддерев.

Індивідуальний контент користувача має міждисциплінарний характер і може мати походження із суміжних предметних областей. У зв'язку з цим виникає задача упорядкування контенту з урахуванням дидактико-семантичних співвідношень між його ділянками. Для розв'язку задачі дидактичного *впорядкування індивідуального контенту* розроблено метод на основі відношень між поняттями онтології із застосуванням апарату стенфордської моделі нечіткого виведення (рис.1.) [4].

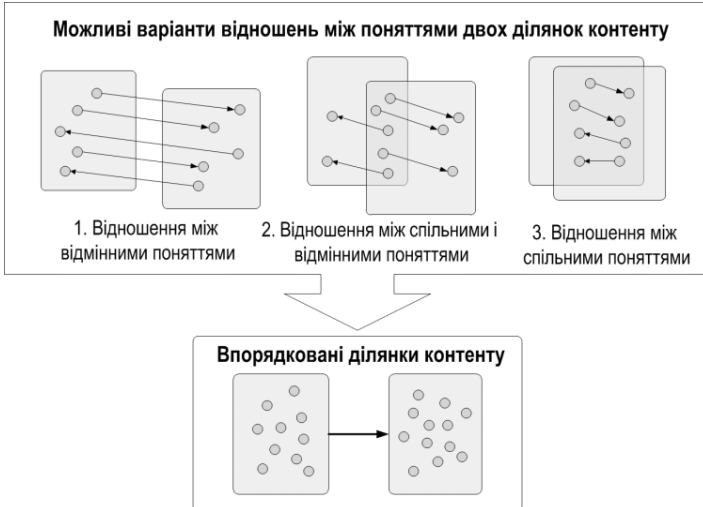


Рис. 1. Схематичне зображення задачі дидактичного впорядкування контенту.

Етапи роботи методу впорядкування індивідуального контенту:

1. Попередній етап побудови транзитивних зв'язків між поняттями і формування транзитивного замикання графу онтології, що ґрунтується на правилі:

$$\text{concept\_before}(c_k, c_l) \langle CF_{kl} \rangle \wedge \text{concept\_before}(c_l, c_m) \langle CF_{lm} \rangle \rightarrow \text{concept\_before}(c_k, c_m) \langle CF_{kl} \times CF_{lm} \rangle$$

Задача вирішується на основі модифікованого алгоритму Флойда-Варшала, де у якості вагів ребер розраховуються фактори впевненості відповідно до поданого вище правила транзитивності.

2. Пошук *цільових* і *фонових* понять кожної з ділянок контенту  $V$  за допомогою правил:

$$\exists t (CT(t) = c \wedge tClass(t) \neq tAttaching \wedge VT(t) \in V) \rightarrow \text{concept\_essential}(c, V)$$

$$\forall t (CT(t) = c \wedge VT(t) \in V \wedge tClass(t) = tAttaching) \rightarrow \text{concept\_pre}(c, V)$$

3. Попарний аналіз відношення слідування між ділянками, на основі правил:

$$\begin{aligned}
 & \text{concept\_essential}(c_k, V_k) \wedge \text{concept\_pre}(c_k, V_l) \rightarrow \\
 & \quad \rightarrow \text{content\_before}(V_k, V_l) \langle CF_{es} \rangle, \\
 & \text{concept\_essential}(c_k, V_k) \wedge c_l \in CVV(V_l) \wedge \\
 & \text{concept\_before}(c_k, c_l) \rightarrow \text{content\_before}(V_k, V_l) \langle CF_g \rangle, \\
 & \text{де } CVV(V) = \{c : VC(c) \cap V \neq \emptyset\}.
 \end{aligned}$$

4. Сортування ділянок контенту за допомогою алгоритму топологічного сортування ациклічного орграфа.

Створене на основі запропонованого формального апарату програмне забезпечення має модульну структуру, зображену на рис.2. Програмний комплекс реалізовано на основі серверної мови PHP, БД MySQL та клієнтських технологій HTML, CSS, JavaScript та Adobe Flex.

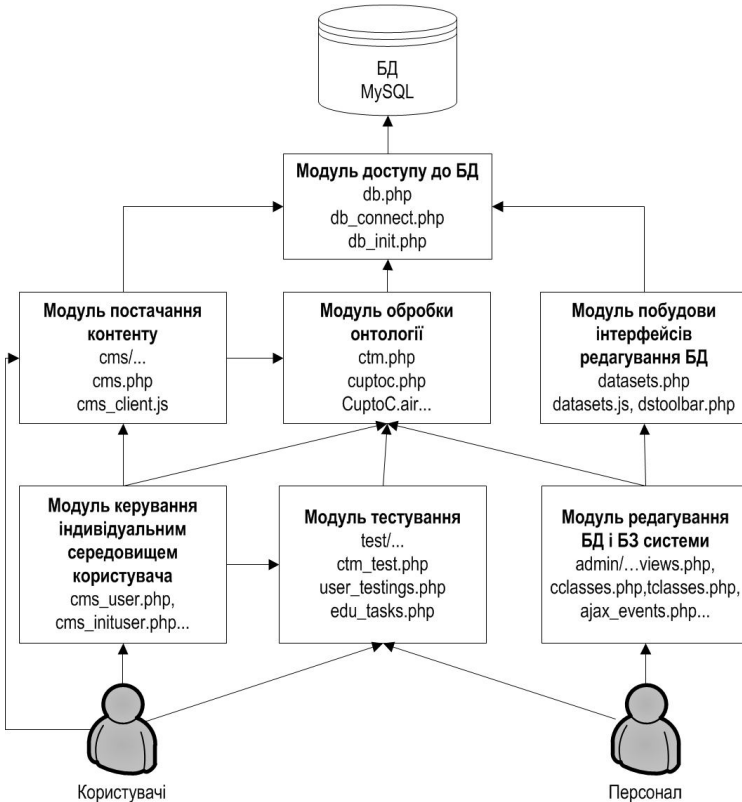


Рис.2. Модульна структура програмного комплексу.

У роботі розроблено комплекс моделей і методів керування інформаційно-навчальним Web-контентом на основі онтологічного підходу, що забезпечує структурно-алгоритмічні основи програмних засобів індивідуалізованого доступу користувачів до затребуваних міждисциплінарних ресурсів інформаційних Web-порталів. Розроблено модель формалізації понятійної складової контенту інформаційно-навчальної Web-системи, що забезпечило основу для методу автоматичної побудови онтології предметної області. Сутність методу полягає в автоматичному визначенні семантичних відношень між структурними елементами формалізованого контенту на базі апарату нечіткого виведення. На основі онтологічного підходу і нечіткої логіки розроблено метод автоматичної побудови індивідуального міждисциплінарного Web-середовища навчання. Розроблено прикладну програмну систему керування контентом з розширеним інструментарієм автоматизованої побудови інформаційно-навчальних Web-порталів, прототип якої функціонує на порталі [znannya.org](http://znannya.org). Поточні дослідження авторів висвітлюються на сайті [www.setlab.net](http://www.setlab.net).

## Література

1. Титенко С. В. Побудова дидактичної онтології на основі аналізу елементів понятійно-тезисної моделі / С. В. Титенко // Наукові вісті НТУУ "КПІ". – 2010. – № 1(69). – С. 82-87.
2. Buchanan B. G., Shortliffe E. H. та ін. Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley, 1984.
3. Титенко С. В. Генерація індивідуального навчального середовища на основі моделі професійних компетенцій у Web-системі безперервного навчання / С. В. Титенко // Вісник Східноукраїнського національного університету імені Володимира Даля – 2009. – №1 (131). Ч.2. – С. 267-273.
4. Титенко С. В. Побудова індивідуального навчального середовища в міждисциплінарному Web-порталі на основі онтології предметної області / С. В. Титенко // XI международная научная

конференция имени Т. А. Таран «Интеллектуальный анализ информации ИАИ-2011», Киев, 17-20 мая 2011 г.: сб. тр./ ред. кол.: гл. ред. С. В. Сирота. □ К.: Просвіта, 2011. – С. 56-61.