

Б. О. КУЗИКОВ, асп. СумДУ, зав. лаб. систем електронного навчання
СумДУ, м. Суми

ВИКОРИСТАННЯ СЕМАНТИЧНИХ МЕРЕЖ І ПОКРИТТЯ ЗМІСТУ З ТЕСТАМИ ДЛЯ ФОРМАЛЬНОЇ ПЕРЕВІРКИ ЕЛЕКТРОННИХ НАВЧАЛЬНИХ МАТЕРІАЛІВ

Зростання обсягів знань вимагає постійну модернізацію навчальних матеріалів, що викликає необхідність впровадження методів контролю повноти та узгодженості множин матеріалів суміжних дисциплін. Стаття пропонує метод формального контролю повноти і логічної впорядкованості викладення навчального матеріалу електронного курсу за допомогою семантичних мереж та покриття контенту питаннями тестового контролю знань.

Рост объемов знаний требует постоянной модернизации учебных материалов. Необходимыми становятся методы контроля полноты и согласованности множеств материалов смежных дисциплин. Статья предлагает метод формального контроля полноты и логической упорядоченности изложения учебного материала с использованием семантических сетей и покрытия контента вопросами тестового контроля знаний.

The rapid growth of knowledge requires an understanding of modernization of educational materials as a sustainable process. It becomes necessary to improve methods of control the completeness and consistency of sets of related e-learning disciplines. The article offers a formal method for monitoring the completeness and logical ordering presentation of educational material through the use of semantic networks and coverage issues of content knowledge test control.

Вступ. Сучасні темпи накопичення знань вимагають від освітніх закладів постійного оновлення навчальних матеріалів. Модернізація існуючих навчальних матеріалів є економічно більш доцільною, у порівнянні із їх повторною розробкою. Часткове оновлення навчальних матеріалів може призвести до неповного охоплення матеріалу, що має вивчатися у курсі, або порушення логіки викладання навчального матеріалу. У рамках одного курсу такі порушення можуть бути виявлені рецензентом – фахівцем у відповідній галузі знань. Виявлення подібних порушень на міждисциплінарному рівні є складним завданням. Крім того, постійний перегляд навчальних матеріалів при ітеративному підході до їх модернізації вимагає додаткових затрат часу експертів.

На теперішній час розроблено ряд критеріїв, що дозволяють оцінити якість дистанційного курсу [1, 2]. Більшість з цих критеріїв можуть бути застосовані лише на рівні окремого навчального об'єкта, як виняток – на рівні курсу в цілому. Дослідження якості курсів на міждисциплінарному рівні, важливість врахування якого підкреслюють ряд дослідників [3, 4] більшістю досліджень не розглядається. Контроль якості курсів при їх модернізації також мало висвітлений в вітчизняній та закордонній літературі [5–8]. Недостатню вивченість даного питання на вітчизняних теренах можна

пояснити метою розробників преш за все “закрити” існуючий брак навчальних матеріалів. Між тим, збільшення кількості навчальних дисциплін, колективна розробка курсів, впровадження розгалужених міждисциплінарних зв'язків вимагає розробки відповідних засобів контролю створення та модернізації курсів.

В СумДУ навчання за дистанційною формою ведеться починаючи із 2002 року. При цьому використовується власна система керування дистанційним навчанням (СКДН СумДУ) [9]. У 2010–2011 навчальному році за дистанційною формою навчаються приблизно 1300 студентів, які вивчають 253 різні навчальні дисципліни за 6 спеціальностями. Стратегія ВУЗу спрямована на постійне підвищення якості навчальних матеріалів. Саме тому при підготовці до 2010–2011 навчального року було модернізовано приблизно 20% навчальних дисциплін. Виконання такого обсягу робіт вимагає застосування різного роду засобів контролю та автоматизації праці зокрема для перевірки повноти та впорядкованості навчальних матеріалів.

Постановка задачі. Метою написання статті є опис підходу застосування семантичних мереж і покриття лекційного матеріалу тестовими питаннями для формального контролю повноти і логічної впорядкованості викладення навчального матеріалу. При цьому під повнотою матеріалу розуміється формальний контроль введення усіх понять, що мають бути вивчені у рамках курсу. Під логічною впорядкованістю матеріалу розуміється наявність порядку вивчення навчального матеріалу (“навчальної траєкторії”) при якому учень може вивчити всі поняття курсу користуючись лише такими навчальними матеріалами, що спираються лише на відомі для нього поняття.

Спосіб перевірки повноти викладення матеріалів. Відомо, що багато проблем знаходять своє рішення в суміжних областях. При розробці складних програмних продуктів важливим є контроль збереження функціональності продукту в цілому при внесенні доповнень в окремі його частини. Модернізація навчальних курсів має подібні проблеми. При внесенні коректив у курс розробники не можуть бути впевнені, яким чином ці зміни впливатимуть на повноту і логічність викладення матеріалу, як у самому курсі, так і на міждисциплінарному рівні. Важливість контролю за регресивною сумісністю збільшується пропорційно кількості та обсягу змін у курсі, обсягу елементарних інформаційних блоків і кількості зв'язків між ними. Особливої гостроти це питання набуває, коли модернізація курсу виконується колективом авторів.

Сукупність теоретичних знань, закладених у лекційному матеріалі можна описати низкою понять, характерних для курсу. Поняття та зв'язки між ними ототожнюються із моделлю знань експерта. Вочевидь, що такий набір понять є досить сталим бо спирається на затверджений паспорт дисципліни. У випадку СКДН СумДУ лекційний матеріал курсу представлено у вигляді гіпертекстових сторінок, ключові поняття вказані лише для частини з них. Побудуємо предметну область першого рівня (за класифікацією

Брусиловського [10]) для подальшого аналізу курсу. Модель предметної області такого типу включає множину понять пов'язаних із текстом, без явного вказання співвідношення між поняттями.

Побудова моделі предметної області курсу є складним завданням, але у ряду випадків може бути частково автоматизована. Поняття курсу c_i – $T_{c_i} \subset T_z \cup T_{q_i} \cup T_{o_i}$ відносяться експертом до однієї з категорій: T_z – множини загальнонаукових понять, що є незмінною, T_{q_i} – множини понять, притаманних лише предметній галузі курсу або $T_{o_i} \subset T_z \cap T_{q_i}$ – множині понять, що мають особливе, відмінне від загальнонаукового, значення в рамках курсу. При побудові предметної області курсу на базі статичного електронного конспекту використовуються припущення щодо скінченності множини понять, логічності послідовності викладення матеріалу в конспектах лекцій, можливості виділити ключові поняття з тексту за допомогою частотного аналізу.

Зв'язок між множиною контенту та концептів у більшості випадків не є очевидним. У рамках моделі, що будується, елемент конспекту – лекція – розбивається на множину фрагментів навчального матеріалу $p_j = \{f_1 \dots f_m\}$. Кожному з навчальних фрагментів відповідає множина пар $F_i = \{\langle t_\phi, s_\phi \rangle, \dots, \langle t_\pi, s_\pi \rangle\}$, де $t_\phi, t_\pi \in T_{c_i}$, $s_\phi, s_\pi \in S$ – множина функціональних станів фрагмента тексту по відношенню до поняття – поняття є базовим для вивчення фрагменту, чи поняття вводиться у ньому. Встановлення відповідності між поняттями та елементами контенту зведено до задачі обрахування міри TF-IDF для відомої множині лексем. Оцінка кількості понять, які необхідні для індексації фрагмента проводиться за допомогою формули Барабаши [12]:

$$N_k = 0.35 + 2.061 \lg(N)$$

Можливість використання математичного апарату безмасштабних мереж доведена у роботі [1]. Так для опису цієї статті достатньо 3–4 ключових слова.

На першому етапі, після того, як для кожної сторінки сформована множина F_i перевіряється умова (1):

$$\forall \{t_j, s_\phi\} \in F_k \rightarrow \exists F_i : \{t_j, s_\pi\} \in F_i, k \neq i \quad (1)$$

— для кожного поняття, вказаного як пререквізит існує інший фрагмент навчального матеріалу, у якому він описаний. Це правило дозволяє контролювати, чи всі поняття застосовані у курсі визначені у ньому. Між тим, а цьому етапі неможливо встановити, чи всі поняття, які мають бути вивчені у

рамках курсу введені у ньому, та, чи не порушується формальний логічний порядок викладення матеріалу.

У більшості випадків для контролю повноти засвоєння матеріалу використовують тести. Це означає, що тестові запитання повинні відображати всю множину знань, запропонований у лекціях, а саме: розглянуті поняття і відносин між ними. Нехай Q_i – множина понять, пов'язаних з питанням q_i .

Зауважимо, що $\bigcup_{i=1..n} Q_i \cap T_{c_i} \neq \emptyset$ – тобто поняття у лекційному матеріалі та

тестових питаннях мають однакову природу. У ідеальному випадку $Q = T_{c_i}$.

Таке співвідношень понять розглянутих у лекційному матеріалі та використаних тестових питаннях дає змогу повністю контролювати стан вивчення курсу. У випадку, коли $Q \setminus T_{c_i} \neq \emptyset$, це позначає, що питання розглядають ширшу множину понять, а ніж описані в лекційному матеріалі, $T_{c_i} \setminus Q \neq \emptyset$ – стан вивчення деяких понять лекції не контролюється. При модернізації курсу розробники повинні намагатися задовольнити умову (2):

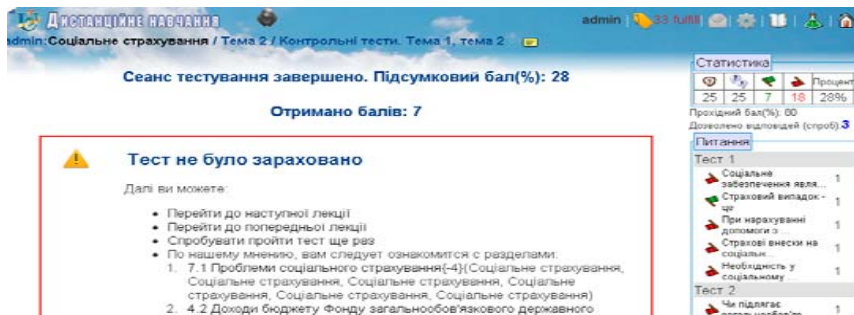
$$Q = T_{c_i} \quad (2)$$

– зміни лекційному матеріалі повинні бути відображені у банку тестових питань і навпаки.

Спосіб використання семантичної мережі для контролю логічної впорядкованості курсу. При модернізації матеріалів курсу можуть виникати дві проблеми – заміна частини матеріалу, таким що несе недостатній обсяг знань, або порушення логічної впорядкованості викладення матеріалу. Повноту викладення матеріалу можна контролювати застосовуючи умови (1) та (2). Другу проблему – формальну логічну впорядкованість викладення матеріалу можна контролювати завдяки аналізу семантичної мережі курсу. Така мережа може бути представлена у вигляді направленого графу, де вузлами будуть елементи контенту, а ребрами – зв'язки між ними. При цьому зв'язки будуються на основі співвідношення понять і елементів навчального матеріалу: кожне ребро йде від елемента контенту, де поняття описане у елемент, де воно відмічено, як таке, що потребує попереднього вивчення.

Ватажимо курс формально логічно впорядкованим, якщо в ньому існує хоча б одна навчальна траєкторія яка складається виключно з прийнятих елементів та при цьому охоплюються всі поняття. Елемент вважається “прийнятним” якщо всі поняття вказані у ньому як пререквізити є вивченими на момент вивчення цього фрагменту. Можна показати, що необхідною і достатньою умовою для формальної логічної упорядкованості курсу є відсутність циклів у графі курсу. Доказ цього положення можна провести використовуючи лему 23.11 та теорему 23.12 [13].

Подальші напрями роботи. В результаті проведених досліджень ми отримали метод формальної перевірки повноти на логічній впорядкованості навчальних матеріалів, що модернізуються. Для цього лекційні матеріали та банк тестових питань були зв'язані між собою через множину понять, які в них використовуються. Отримана таким чином семантична мережа є моделлю знань експерта. Результати тестового контролю знань – відображають знань студента пов'язані з предметною областю курсу. Використовуючи результати тестування можна виділити фрагменти теоретичного матеріалу, пов'язані із відповідями студента на кожне з питань тестового контролю. Приклад реалізації такого підходу в СКДН СумДУ показано на рисунку 1.



Адаптивна навігація, на основі покриття контенту тестовими питаннями

Використовується наступна модель для визначення навчальних матеріалів, що потребують додаткового вивчення:

$$V_k = M \times \frac{aG' \times Q - C \times R_k}{a} - M \times \frac{C \times R_k - bG' \times Q}{b},$$

де C – вектор складності тестових питань, R_k – вектор результатів тестування, $g_i \in G'$ – ймовірність того, що поняття $t_i \in G$ вивчене повністю. Значущість попередньої оцінки стану вивченості поняття коригуються за допомогою коефіцієнтів $a, b \in [0..1]$. Щоб встановити відповідність між концептами та тестовими питаннями використовується матриця зв'язності

елементів навчального контенту $M = \|f_{i,j}\|$, де $f_{i,j} = \frac{|Q_i \cap F'_{j_{pr1}}|}{|F'_j|}$, Q_i – множина

понять, пов'язаних з питанням q_i , $F'_{j_{pr1}}$ – множина базових понять для фрагменту f_j . Матриця M аналогічна до матриці зв'язності, але замість множини результуючих понять $F'_{j_{pr1}}$ використовується $F''_{j_{pr1}}$ – множина

базових понять [14]. Підсумовуючи вище сказане, відмітимо, що подальшими напрямками роботи будуть вдосконалення методів побудови предметної області курсу на базі конспектів лекцій, ґрунтовне дослідження ефективності запропонованого підходу до побудови адаптивної навігації та вдосконалення засобів для побудови більш тісної міждисциплінарної інтеграції навчального процесу.

Висновки. У статті наведено метод, що дозволяє сформувати систему правил, перевірка яких дозволяє здійснювати формальний контроль повноти і впорядкованості курсу. Застосування методу дозволяє зменшити витрати часу на рецензування курсу при ітеративному підході до його модернізації. Підкреслимо, що підхід дозволяє проводити контроль на рівні курсу та міждисциплінарному рівні. Крім того побудова семантичної мережі курсу дозволяє запровадити адаптивне керування навчальною діяльністю студента, що є додатковою перевагою. Зазначений метод було впроваджено у СумДУ для контролю якості модернізації навчальних дисциплін за дистанційною формою навчання.

Список літератури: 1. Гаврилов Н.А. Некоторые аспекты анализа эффективности формирования дистанционных образовательных ресурсов // Дистанционное и виртуальное обучение. – 2010. – № 2. – С. 15–22. 2. Соколов А.С. Деякі аспекти синтезу комп'ютеризованої адаптивної системи навчання // Автоматика. Автоматизация. Электротехнические комплексы и системы. – 2008. – № 2. – С. 89–93. 3. Мазурок Т.Л. Ассоциативный подход к моделированию системы межпредметных связей а АОС // Сборник научных трудов IV семинара «Информационные системы и технологии». – Одесса, ОГАХ, 2006. – С. 155–162. 4. Мазурок Т.Л. Интеллектуальное управление обучением в виртуальной среде // Новый колеріум. – 2008. – № 6. – С. 48–55. 5. Babič F., Wagner J. Trialological learning – new approach to education. // 5th International Conference on Emerging e-learning Technologies and Applications ICETA. – Stara Lesna, Slovakia, 2007. – С. 183–187. 6. Оганесян Е.В. Метод динамического тестирования в определении структуры знаний студентов // Alma mater. – 2008. – № 9. – С. 43–46. 7. Шишов Е.В., Ундозерова А.Н. Комплексный подход к разработке средств учебного назначения для e-learning при подготовке инженеров–программистов в высшей школе // Дистанционное и виртуальное обучение. – 2009. – № 5. – С. 27–37. 8. Leading the Learning Revolution // Individual Training & Education Modernization Update. – 2009. – №1. 9. Vasylyev A., Lavrik T., Lyubchak V. System of distance education at Sumy State University // 5th International Conference on Emerging e-learning Technologies and Applications. – Stara Lesna, Slovakia. – 2007. – С. 291–294. 10. Brusilovsky P. Adaptive hypermedia, an attempt to analyze and generalize // Lect.notes. Comput. Sci. – 1996. – № 1077. – С. 288–304. 11. Salton G. Theory of Indexing // Society for Industrial and Applied Mathematics. – Philadelphia. – 1975. – 56 c. 12. Barabasi A, Reka A. Emergence of scaling in random networks // Science. –1999. – №15. С. 509–512. 13. Cormen T., Leiserson C., u dp. Introduction to Algorithms. – MIT Press. – 2001. – 1202 c. 14. Kuzikov B.O. Creation connectivity matrix of E-content elements for distance learning // Nauka i studia. – 2009. – № 6(18). – С. 62–66. 15. Lubchack V., Kuzikov B., Kirichenko K. Using semantic web and covering context by test for course formal testing // 8th Int. Conference on Emerging eLearning Technologies and Applications, High Tatras, Slovakia. – 2010. – С. 135–140.

Надійшла до редколегії 05.11.2010