

АВТОМАТИЗИРАНО ГЕНЕРИРАНЕ НА ТЕСТОВИ ВЪПРОСИ, ОЦЕНЯВАЩИ ЗНАНИЯ ПО ТАКСОНОМИЯТА НА БЛУМ

Марияна Соколова, Христо Инджов, Георги Тотков

*mariana_sokolova@yahoo.com, h.indzhov@pastelstudios.com,
totkov@uni-plovdiv.bg*

Резюме: Съвременен направление в развитието на системите за е-обучение е адаптиране към когнитивните способности и знания на обучаваните. При електронното тестване, например, вече се говори не само за набор от стандартни тестови въпроси (напр. с множествен избор), но и за тестове, измерващи равнището на знания (напр. по таксономията на Блум). В работата се предлага и експериментира методика - как оценени знания на обучавани (в дадена предметна област) да се използват за конструиране на следващи тестови въпроси. Методиката е общоприложима (не зависи от областта, дисциплината, обучаваните и др.) и адаптивна (вкл. с автоматизирано генериране на тестови въпроси относно понятия и отношения в конкретната предметна област).

Ключови думи: разширена таксономия на Блум, електронно тестване, автоматизирано генериране на тест, адаптивен тест,

ВЪВЕДЕНИЕ

Оценяването на знания на обучаваните е важна част от учебния процес. Най-използваният начин за оценяване в съвременните системи за е-обучение (СеО), е т.нар. електронно тестово изпитване (е-тестване), което определя и големият интерес на изследователи, методисти и педагози към е-тест.

СеО предоставят множество средства, които подпомагат и развиват процеса на е-тестване. Изследват се и различни нови методики и средства за е-тестване (концептуални карти [1], мисловни карти [12], портфолио [2], 360 градусова обратна връзка [8], адаптивни тестове [5, 10, 17], автоматично генерирани тестове [24, 25], тестове за измерване на знание от по-високо ниво [14, 15], акумулативни тестове [14, 15] и др.). Електронният тест, предлаган на обучаваните, в крайна сметка може да се представи като линейна или нелинейна ([5, 10, 17]) редица от тестови въпроси (от различен тип).

Основни проблеми при тестването са надеждност, валидност и коректност (на теста, тестовите въпроси и оценяването). За да се преодолеят проблемите, тестовете трябва да се съставят според ясни критерии и цели, а включените задачи трябва да съответстват на равнището на знания на обучаваните. До неотдавна се считаше, че за измерване на равнището на знания от по-високо ниво в дадена предметна област (ПО) могат да се използват само въпроси от тип 'есе'. Напоследък се коментира и обосновава възможността - въпроси от тип 'избор' да измерват знание от по-високо ниво от таксономията на Блум [7]. Тестовите въпроси могат да бъдат подбирани и подреждани по различни критерии – трудност, тип, време за решаване, съответна предметна област и понятия, възможност (или не) за автоматично оценяване и т.н. Други, често използвани критерии при съставянето, са – въпросите да са подходящи за по-големи групи обучавани и способност за измерване на различни равнища на знанието. През 1956 г. Блум [3] предлага 6 (шест) равнища за оценяване - знание, разбиране, приложение, анализ, синтез, оценяване. По-късно тази таксономия е развита и представена като двумерна рамка, при която по първата ос са разположени 6-те равнища на знанието (именувани тук запамятаване, разбиране, прилагане, анализ, оценяване и създаване), а по втората ос са разположени различните типове когнитивни процеси, засегнати при различните равнища на знанието (факти, понятия, процедури, метазнание [6]). Таксономията на Блум се използва в множество СеО [4, 11, 13, 18], с което се цели измерване на по-сложни знания, постигане на по-високо качество на обучението, прилагане на различни педагогически методи и схеми.

В работата се предлага и експериментира методика - как знания на оценявани обучавани (в дадена ПО) да се използват за конструиране на следващи тестови въпроси (в същата ПО). Методиката е общоприложима (не зависи от областта, дисциплината, обучаваните и др.) и адаптивна (вкл. с автоматично генериране на тестови въпроси относно понятия и отношения в конкретната ПО).

МЕТОДИКА ЗА ГЕНЕРИРАНЕ НА ТЕСТОВИ ВЪПРОСИ

При съставянето на задачите, включени в е-тест, са възможни няколко подхода: а) авторски ('ръчно' създаване от експерти в ПО); б) автоматично генериране на базата на семантично описание на ПО; в) използване на техники от областта на компютърната лингвистика - извличане на данни от текст и/или Интернет-източници, включване на 'въпросно-ответни системи' и др.; г) избор измежду стандартен набор на шаблонни въпроси (с последващо редактиране за съответната ПО); д) автоматично генериране на базата на предишен опит при е-тестване, и др.

В работата се разглеждат конкретни решения, свързани с г) и д). На базата на разширената таксономия на Блум е създадена таблица (ще бъде представена в следваща публикация), съдържаща

стотици въпроси и шаблони на въпроси, независещи от изучавана ПО, дисциплина и тема. Таблицата се състои от 4 (четири) колони, представлящи когнитивните умения, свързани съответно с факти, понятия, процедури и метазнание, и 6 (шест) реда за равнищата на измерваното знание – запаметяване, разбиране, прилагане, анализ, оценяване и създаване. Във всяка от 24-те клетки на таблицата са дадени конкретни тестови въпроси и/или шаблони на въпроси – представители на измерваното равнище на знание (според реда, в който е разположена съответната клетка) и на усвоените когнитивни умения (според съответния стълб). Например, представители на въпроси, свързани с измерване на равнище ‘запаметяване’ относно когнитивно умение ‘усвояване на понятия’ са: *Избройте основни факти, идеи, понятия, процедури в темата?; Опишете; Посочете примери за; Дайте определения за; Избройте основните отношения между* и др.

Таблицата е силно генеративна – отговори, дадени от обучаваните на въпроси/шаблони на предходно равнище на оценяване (например в даден ред/стълб) се обработват и запаметяват, с цел следващо използване като източник на конкретни следващи въпроси (с попълване на шаблони в друг ред/стълб).

С дирижиран избор на равнище на знание (в ред) и когнитивно умение (в стълб), автоматизирано могат да се конструират е-тестове с избор на въпроси/шаблони от съответната клетка (сечение на избраните ред и стълб). Според добрите педагогически практики е добре въпросите в е-тест да се подреждат от по-ниско към по-високо равнище, с постепенно увеличаване на трудността (вкл. и автоматично определяна), и да представят всички равнища на знания и когнитивни умения според обновената таксономия на Блум.

Важно е да се отбележи, че въпросите в разглежданата таблица са от т. нар. ‘акумулативен тип’ [16], и могат безпроблемно да бъдат ‘вградени’ в автоматизирана система за е-тестване. Акумулативното генериране на въпроси позволява генериране на нови тестови въпроси от различен тип (на базата на отговори, натрупани в процеса на е-тестване). По същество, таблицата представя иновативна методика, широко приложима не само при създаване на тестови бази от данни и/или тестови задачи към отделни учебни теми, но и за проектиране и създаване на системи за е-тестване с широко приложение.

ПРИМЕР: ГЕНЕРИРАНЕ НА ТЕСТОВИ ВЪПРОСИ ПО БЛУМ

Въпрос от **първо (най-ниско равнище)** за измерване на запаметяването на информация при автоматично генериран тест, например, е *„Избройте 5-те най-важни понятия - предмет на разглежданата тема“*. Възможните акумулирани (евентуално оценени и редактирани) отговори на група обучавани да означим с $P = \{O_1, O_2, \dots, O_p\}$, а със C – *подмножеството на коректните (положително оценени) отговори от P*. Тогава, на следващи оценявания може да се предлага друг въпрос (с използване на вече оценени отговори от P), например – *„Дайте определение за понятие O_i “* (i се избира случайно, с единствено условие $O_i \in C$). Възможно е и допълнително параметризиране – например по брой изучавани понятия (в примера – 5) и отношения, и т.н.

За измерване на **второто равнище ‘разбиране’**, е възможно генериране на въпрос *„Посочете отношението между понятия O_i и O_j “* (тук $O_i, O_j \in C$, с варианти - множествен избор от списък на универсални отношения или създаване на списък от оценяваните. За акумулиране на отношения в ПО могат да бъдат генерирани и зададени множество въпроси (поне $\frac{n(n-1)}{2}$, където $n = |C|$). След натрупване на отговори на обучаваните в базата данни ще разполагаме с множество $R = \{R_1, R_2, \dots, R_r\}$ от възможни отношения между понятия в дадената ПО. Аналогично, елементи на R могат да бъдат използвани за генериране на серия от следващи въпроси. Друг въпрос от същото равнище на трудност е *„Подрежете понятия $O_1, O_2, O_3, \dots, O_n$ (от C) по групи и категории“*, като групите и категориите се предлагат под формата на списъци (за факти, имена, събития, действия, дати, личности, места, фрази, определения, принципи, теории, методи, процедури и т.н.), към които може да се добавя. Възможно е SeO да се базира и на съответна онтология (предмет на друго разглеждане).

Въпроси от равнище **‘измерване на умения за анализ’** са *„Сравнете понятия O_i и O_j “* и *„Определете типа на връзка между понятия O_i и O_j “* (отново може да се избира от акумулирания списък R_1, R_2, \dots, R_n и/или от стандартен списък на отношения, например – принадлежи_на, част_от, и др.

Примерни въпроси, подходящи за по-високи равнища на знанието, които могат да се генерират (на базата на отговори, получени при оценяване на по-ниски равнища) са: *„Подрежете понятия O_1, O_2, \dots, O_n по значимост за темата“ (равнище ‘оценяване’); „Кое от понятия O_1, O_2, \dots, O_n е централно за темата, и защо?“, „Разкажете със свои думи за (централното) понятие O_i (‘създаване’), и т.н.*

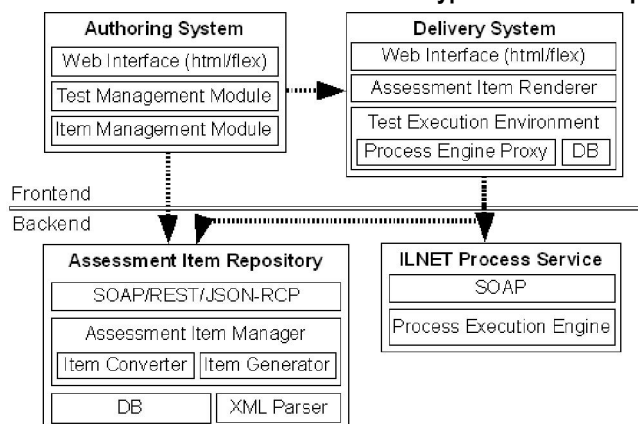
ПРОТОТИП НА КОГНИТИВНА ТЕСТОВА СИСТЕМА

В резултат на проведеното изследване относно автоматизирано генериране на въпроси за оценяване на по-високи равнища на знанието е създаден прототип на система за е-тестване (*ProtoSeT*). Прототипът се състои от **четири подсистеми** за: създаване на тестове, доставяне и визуализиране на тестови задания, управление на хранилище за тестови въпроси и сервиз за изпълнение и управление на тестови процеси. Първите две подсистеми обслужват клиентската част (за потребители – обучавачи и обучавани)

на *ProtoSeT*, а останалите (изцяло сервизно ориентирани) - свързвателната част. Фиг. 1. представя детайли на системната архитектура. Комуникацията между модулите се осъществява с помощта на протокола SOAP [19], с което се осигурява платформена независимост и преносимост.

Подсистемата за **създаване на тестове** се състои от два модула: Мениджър на тестови въпроси (за добавяне, редактиране, натрупване и изтриване на въпроси и акумулирани отговори) и Мениджър на тестове (за моделиране и съхраняване на процеси на оценяване под формата на работни потоци).

Фигура 1. Системна архитектура на *ProtoSet*



Подсистемата за **доставяне и визуализиране на тестове** координира изпълнението на тестовите модели и акумулира (в хранилището) отговорите на обучаваните на отделните тестови въпроси. Визуализацията на отделните тестови въпроси, както и начинът им за представяне (в рамките на съответния тест), също е част от функционалността на тази подсистема. Режимът за *визуализация* може да се адаптира към различни категории обучавани (позволява един и същи въпрос да бъде визуализиран по различни начини, вкл. в зависимост от специални образователни потребности).

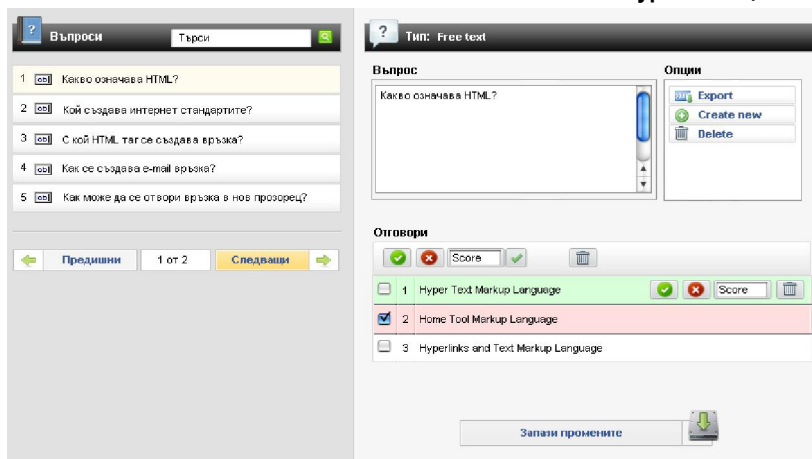
Подсистемата за **управление на хранилището на тестови въпроси**, включва модул за управление (мениджър), трансформатор и генератор на тестове, и е базиран на съответна релационна база данни. Достъпът до подсистемата може да бъде осъществен по протоколи SOAP, REST и JSON-RCP. Възможността за използване на протокол JSON-RCP предразполага към използване на хранилището за тестови въпроси директно от браузър, и като следствие – за изграждане на модерни Интернет-приложения. Мениджърът предоставя средства за добавяне, изтриване и редактиране на въпроси, като в процеса на тестване акумулира отговори и нови въпроси. Трансформаторът преобразува акумулиращи тестови въпроси в/от стандартни (от различен тип – за множествен избор, с посочване, и др.). Генераторът на тестови въпроси е отговорен за създаването на въпроси по дадени критерии като равнище на трудност, резултат от предишен въпрос и т.н.

Сервизът за управление на тестови процеси се базира на т.нар. ILNET-система за моделиране и изпълнение на работни потоци [9]. Всеки учебен тест се моделира като работен поток под формата на граф (от тип мрежа на Петри), изпълняван и управляван от ILNET-сервиза.

Всеки акумулиращ тестов въпрос 'събира и съхранява' отговори на обучавани, получени по време на електронното оценяване. Както беше посочено, така акумулираните данни са в основата на следващо генериране на стандартни типове тестови въпроси (за *ProtoSeT* - множествен избор, свободен отговор, множествен отговор, въпрос от файл и текстов отговор, отговор от файл, въпрос и отговор от файл, множествен избор и отговор от картинки).

За реализация на *ProtoSeT* (Фиг. 2.) е използвана софтуерната рамка на Zend [20] (среда за разработка Zend Studio), известна с високото си качество, надеждност и голяма общност от поддръжници. Zend осигурява стабилна основа за следващо развитие на прототипа, използва широко разпространен език за Интернет-програмиране (PHP) и позволява бърза и безпроблемна интеграция с други CeO.

Фигура 2. Общ изглед на ProtoSeT



ЗАКЛЮЧЕНИЕ

Проведени са експерименти по съставяне на адаптивни тестове с използване на методики за измерване на всички нива на знанието по разширената таксономия на Блум (от 'запомняване' до 'създаване'). На базата на акумулирани отговори на въпросите от теста са генерирани стандартни тестови въпроси.

Автоматизирано генерираният тест в *ProtoSeT* не зависи от конкретната ПО, но се конструира на базата на степента на усвояване/разбиране (на съответната тема) от обучаваните. В този смисъл, *ProtoSeT* може да се определи като **адаптивна** (следващ въпрос се генерира на базата на отговори, дадени на предходни въпроси от конкретната група обучавани) и **акумулативна** (отговорите се натрупват и се използват за съставяне на тестови въпроси от различен тип - множествен избор, истина/лъжа, попълване на празнини и др., за следващи групи оценявани обучавани) **тестова система**.

Реализирането на подсистема за акумулиращи въпроси (от програмистка гледна точка) е сравнително лесно, но за поддържане и на стандарти (напр. IMS) е необходимо специално разширение [14].

Планирано е следващо усъвършенстване на *ProtoSeT* с цел: автоматично генериране на концептуални карти на изучаваната ПО (вкл. с добавяне на фокус-въпроси); автоматично оценяване на знания и умения в ПО (спрямо експертна концептуална карта); интегриране на онтологии за класификация на изучаваните факти, понятия, отношения, процедури и др.

Предстои интегриране на *ProtoSeT* със CeO, и експериментиране на методика за автоматизирано създаване на хранилища от тестови задачи в различни ПО. Основната идея тук е - след всяка 'изучена' тема в съответната ПО да се генерира тест от въпроси, свързани с различни равнища на оценяваното знание и когнитивни умения (съгласно представената методика).

Работата е частично финасирана от проекти МИ-203 и Д002-308 към Националния фонд за научни изследвания.

ЛИТЕРАТУРА

- [1] Anohina, A., D. Pozdnakovs, J. Grundspenkis. Changing the Degree of Task Difficulty in Concept Map Based Assessment System. Proceedings of the IADIS International Conference "e-Learning 2007", July 6-8, 2007, Lisbon, Portugal, pp. 443-450.
- [2] Barrett, H. (2004). Differentiating Electronic Portfolios and Online Assessment Management Systems. In R. Ferdig et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2004*, pp. 46-50.
- [3] Bloom, B. S. *Taxonomy of Educational Objectives* (1956). Published by Allyn and Bacon, Boston, MA. Copyright (c) 1984 by Pearson Education.
- [4] Chen S., R. Kuo, M. Chang, T. Liu, J. Heh. Developing True/False Test Sheet Generating System with Diagnosing Basic Cognitive Ability. In the Proceedings of the AACE 20th World Conf. on Educational Multimedia, Hypermedia & Telecommunications, Vienna, Austria, June 30-July 4, 2008, 5740-5748.
- [5] Conejo R., Ed. Guzmán, E. Millán, M. Trella, J. L. Pérez-De-La-Cruz, A. Ríos, SIETTE: A Web-Based Tool for Adaptive Testing, *International Journal of Artificial Intelligence in Education*, 14 (2004) 1-33.
- [6] Dalton E., The "New Bloom's Taxonomy," Objectives, and Assessments, by December 3, 2003.
- [7] DAMMCQs: Appendix. C: MCQs and Bloom's Taxonomy, <http://web.uct.ac.za/projects/cbe/mcqman/mcqappc.html#C2>.
- [8] Focal360, 360 Degree Feedback Survey System - 360 Evaluation and Assessment Software for Feedback Surveys <http://www.custominsight.com/360-degree-feedback/>.

- [9] Indzhov Hr., D. Blagoev, G. Totkov, Executable Petri Nets: Towards Modelling and Management of e-Learning Processes, ACM International Conference Proceeding Series; Vol. 375, Proc. of the 10th Int. Conf. on CompSysTech, Rousse, Bulgaria, June 18-19, 2009. IIIA.12.1 - IIIA.12.6.
- [10] Kingsbury G., Zara, A. R., Procedures for Selecting Items for Computerized Adaptive Tests. Applied Measurement in Education, 2(4), 1989, 359-375.
- [11] Muzio J. A., T. Heins, R. Mundell, CEDAR Experiences with Reusable e-Learning Objects: From Theory to Practice The Internet and Higher Education Volume 5, Issue 1, 1st Quarter 2002, 21-34.
- [12] Online Mind Mapping and Brainstorming – MindMeister, <http://www.mindmeister.com/>.
- [13] Peiwen T., E. T. Yew, Teaching Thinking Skills in e-Learning – Application of the Bloom's Taxonomy, Frontiers in Artificial Intelligence and Applications; Proceeding of the 2005, Vol. 133, 678-681.
- [14] Sokolova M., G. Totkov, Extended IMS Specification for Accumulative Test System, ACM International Conference Proceeding Series; Vol. 374, Proc. of the 9th Int. Conf. on CompSysTech, Gabrovo, Bulgaria, June 12-13, 2008, V.14-1–V.14-6.
- [15] Sokolova M., Totkov G., Accumulative Question Types in e-Learning Environment, International Conference on Computer Systems and Technologies - CompSysTech'2007, IV.21-1 - IV.21-6.
- [16] Sokolova, M., G. Totkov, Towards accumulative e-learning systems. 3rd Balkan Conference in Informatics, 2007, Sofia, I.513 - I.522.
- [17] Totkov G., E. Somova, M. Sokolova, Modelling of e-Learning Processes: an Approach Used in Plovdiv e-University, Int. Conf. on Computer Systems and Technologies (e-learning), 17-18 June 2004, Rousse, IV.12.1 – IV.12.6.
- [18] Wikramanayake G.N. e-Learning: Changes in Teaching and Learning Styles. Proceedings of 22nd National Information Technology Conference (NITC) on e-Sri Lanka from Vision to Reality, 2003, July 3-4, 118-124.
- [19] W3 SOAP Specification, <http://www.w3.org/TR/soap12-part1/>.
- [20] Zend Framework, <http://framework.zend.com/>.