Enhanced test evaluation for web based adaptive learning paths

Daniel Szentes, Bela-Andreas Bargel, Alexander Streicher, Wolfgang Roller
Department Interoperability and Assistance Systems
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB
Karlsruhe, Germany

Abstract—**In our work we describe an approach for the automatic evaluation of student test results by a web based e-learning application, which considers more criteria than the obvious correctness. This evaluation is focused on the use in adaptive learning paths, meaning that the student's learning path is adapted based on the test results. We describe a concept which integrates the evaluation of more criteria than correctness and how an implementation of such adaptive learning paths could look like in a web based e-learning environment. Our work shows a working prototype which uses the enhanced criteria speed and certainty based upon an evaluation with fuzzy logic algorithms. Finally we present some first results about the acceptance of web based adaptive learning paths.**

Keywords-Adaptive e-learning, adaptive learning paths, test evaluation, web based application, fuzzy logic

I. Introduction

Many e-learning programs follow the linear model of the behaviorist Burrhus Frederic Skinner and his thoughts for a teaching machine in the 1950s [1]. The learner is forced to complete all learning units in a predefined linear order, or to speak in a schoolbook metaphor, he is supposed to read and learn all chapters in a rigid order. A successful test qualifies the learner to go on with the next unit or chapter.

Norman A. Crowder's intrinsic programming [2], developed in the same decade as Skinner's teaching machine, can be seen as the basic idea for adaptive learning paths. Crowder also presented a test to the student (or more precise a multiple choice question) after a learning unit. But when the test was unsuccessful, the student in Crowder's model is transferred to a special chapter where the cause of his error is explained and then is redirected back to the test. Crowder wanted to create an automatic tutor which offers individualized help based on the learner's performance. His model is also called branched programming, showing that this is the basic idea of redirecting a learner to another path according to the individual learning success. This offers the possibility for a learner to reach educational objectives more effectively.

Crowders branched programming is to some degree the ancestor of modern adaptive learning path systems. These modern systems may only provide some more "types of branches" or types of components of the course structure, leading to a higher-order model [3]. Fig. 1 shows a sample adaptive course structure. It is presented as a graph where the nodes are the learning units (LU) and the LUs are linked with edges. It shows the components branch (A), confluence (B), iteration (C) and elective sequence[1] (D).

When an author creates such a course graph, there is a fair amount of possibilities which path the student could follow, while working through the learning units. The question is, how the e-learning application can automatically "know" which path to follow. The classic solution for this is to apply and evaluate tests, like Skinner and Crowder did within their teaching machines. Crowder used a multiple choice question as test and focused on the correctness of the given answers [2]. In this paper we like to discuss if it is reasonable to enhance the evaluation with further criteria than just correctness and how to implement such a concept for a web based e-learning application.

II. Enhanced Test Evaluation for Web Based Adaptive Learning Paths

The obvious criterion to assess the quality of test results is to check which answers were correct and which answers were wrong. In our opinion this is not sufficient enough for a fair evaluation of a test and we would like to point out the following thought:

Should you grade a multiple choice test better if

- *3 of 5 answers were correct within 90 seconds, or if*
- *5 of 5 answers were correct in third attempt, or if*
- *a pupil from elementary school has 1 of 5 correct answers though the question was designed for academic students?*

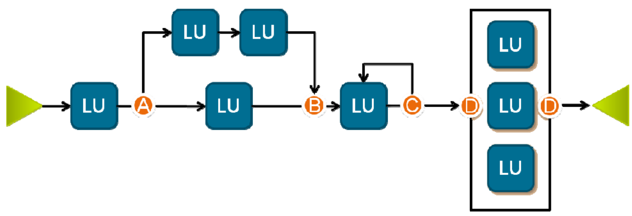

Figure 1. Sample course path with branch A, confluence B, iteration C and elective sequence[1] D

[1] Elective sequence means, that the order of the sequence is facultative, but all learning units must be completed.

As this question indicates we should consider more than the criterion of correctness. A teacher would adapt to a “slow” student even if all his answers were correct. In this case he might present him a different path where the student could understand the material better. The question above not only shows indicators for the criterion speed, but also for the criteria certainty and difficulty level. These criteria seem to be of relevance for an author in an e-learning environment in order to enable more individualization.

A. Concept

The general idea is to integrate the criteria speed, certainty and difficulty level into the classic evaluation of a test when the e-learning application calculates which path the student should follow. The criterion correctness still remains the basic criterion.

Correctness means in our concept that we regard the amount of correct answers in an exercise in comparison to the amount of incorrect answers – if there are any (this depends on the exercise type). This yields a preliminary grade for the exercise.

The difficulty level is supposed to take into account that an exercise can vary in difficulty according to different target groups. This means that an author of the learning unit and of the resulting test should consider whether he has a homogeneous target group in terms of previous knowledge levels, or not. If there is an inhomogeneous target group, we think that it is an advantage to respect the different levels and to provide an individual evaluation of the exercise. The base grade of the correctness should be altered according to the level of knowledge. So when the pupil from the elementary school (see example above) does the exercise it will get a positive adjustment to the result, whereas the professor as domain expert might get a negative adjustment to the result. In this case we ignore that the correct answer might just be guessed, but we are aware of this possibility. However, in order to find an adequate adjustment the system should provide some kind of competence model for the learner. A competence model for an e-learning system tries to mirror the actual knowledge level of a person. Competence in general encompasses a combination of potential knowledge, skills and behavior to properly perform a specific task. In an e-learning application such a task could be to solve a test which requires a specific competence level. An e-learning application typically focuses on the user’s knowledge to model the competency. This competence model is the basis for the suggestion which learning path the user is supposed to take because it reflects how much the user most probably knows about a certain topic. By following a certain learning path a user acquires certain knowledge, i.e. the user learns, and therefore the competence model can be updated to reflect that new learning state. Tests and exams can check if the user has actually learned the very knowledge of the learning path. The e-learning application itself can be made adaptive, i.e. it can learn to adjust its model of the learner. The internal state of the application can be seen as the student model, or user model, in which the competence of the user is encoded. How to effectively represent, learn or adapt user models [4], [5] and how to model competency for learning systems, e.g. for teaching models and student models [6], is a topic of great interest in the scientific learning community ([7], [8], [9]) and won’t be further detailed in this paper.

The criteria speed and certainty are at least implicit part of every test analysis. Speed is assessed by the time from the beginning of an exercise until hitting the ‘evaluate’-button. For being able to measure the beginning the learner also has to ‘open’ each exercise. Regarding the expenditure of time an author has to consider the complexity of each exercise. The time may be used to differentiate between test aspects of knowledge level on the one hand and speed on the other [10].

Certainty is counted as number of attempts while answering the exercise. A threshold set by the author defines a maximum of iterations. Important aspects of the test criterion certainty involve the statistical probability of guessing the solution as mentioned before. Multiple-choice and especially single-choice exercises with only a few answering options are subject for this. But also “test-wiseness” [11] plays an important role and should be considered. Nevertheless a criterion like certainty might be interesting when using exercise types like cloze. A second attempt of a nearly correct answer might be useful for the learner to correct spelling mistakes or even synonymous answers. However, the e-learning system itself might take this into account by using adequate error-tolerant algorithms like the Levenshtein edit distance or semantic extensions like a synonym web service [12].

In a mathematical sense the criteria speed and certainty seem to be “weak”, if terms like “this answer was sure” or “this answer was only quite sure” has to be interpreted by evaluation logic in a web based e-learning application. To combine and represent these “weak” criteria in a fair way, classic evaluations like the so called *target model* [13] are not sufficient. We decided to develop an evaluation logic using the ideas of the fuzzy set theory proposed by Lotfi Asker Zadeh 1965 [14]. Our concept proposes fuzzy sets on linguistic variables for the criteria correctness, speed, certainty and difficulty level. By applying truth values derived from the actual values of a student and a subsequent defuzzification a numerical result value can be calculated. This means that not every criterion has to be part of the equation and it allows even more criteria to be added into the equation.

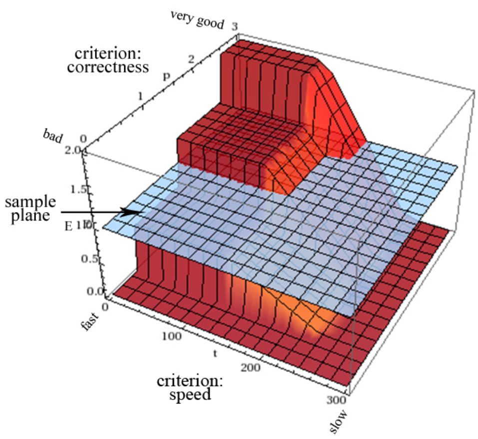

Figure 2. characteristic diagram for the criteria speed, correctness and result value E (t = expenditure of time, p = right/wrong quota).

Fig. 2 shows a characteristic diagram for the criteria speed and correctness. The plane based upon the result value E serves as decision criterion which path to follow next in the e-learning course. For further mathematical details of our proposed evaluation logic see [15].

B. *Implementation of concept*

The prototype of the adaptive learning path was implemented in the web based e-learning and authoring environment Crayons[2]. Due to the complexity of a competence model our first implementation of the concept passes on the criterion of difficulty level and puts the criteria correctness, speed and certainty into effect. The evaluation logic for the exercises has been implemented as described above with the mentioned criteria. The result value of the fuzzy logic evaluation can be seen as a grade for the student. When there is a decision point in a learning path a test consisting of one or more such exercises decides which path to follow next. The arithmetic average over all evaluated exercises of the test leads to the final result of the test and the decision which path to follow. Fig. 3 shows the simplified workflow and Fig. 4 an example decision point at a branch in the e-learning course.

Because of the complexity of the evaluation process several problems are to be solved in order to get an intuitive web based front end and presentation for both authors and students.

So the authoring environment compensates the complexity of the fuzzy based evaluation logic and the authors need to quickly create good quality content. The authoring front end provides a layer of abstraction to the underlying complex evaluation logic. Fig. 5 shows the user front end for setting the parameters for the criteria. In the shown example the test consists of three exercises, a multiple choice, a cloze and an allocation exercise. A slider represents three areas for the criteria certainty and speed. A student will have a positive adjustment (improvement) to his grade in the area between 0 and the first parameter on the slider. The third parameter at the end of the slider shows the maximum of iterations or time. The area between parameter 2 and the maximum defines an area where the student will get a negative adjustment (degradation) on his grade. Leaping the maximum means, that the exercise is automatically rated with the worst grade. The grading according to correctness can also be altered via slider. In order to link the grades to learning units, the author simply adds a grade and chooses a destination out of a list of the existing learning units (see Fig. 6).

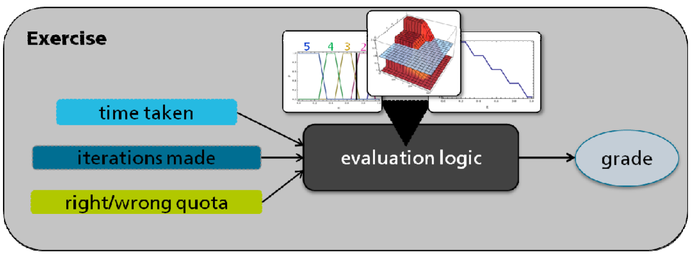

Figure 3. Simplified workflow of exercise evaluation.

[2] Crayons® is a platform independent and SCORM compatible e-learning and authoring environment (http://crayons.fraunhofer.de).

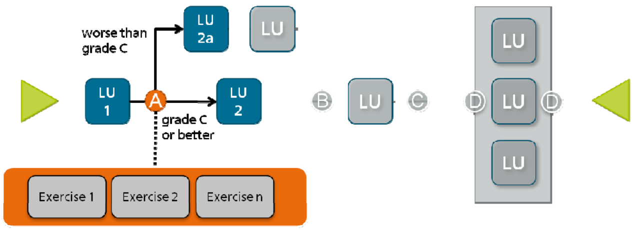

Figure 4. Test at branch A with decision criteria for successor path.

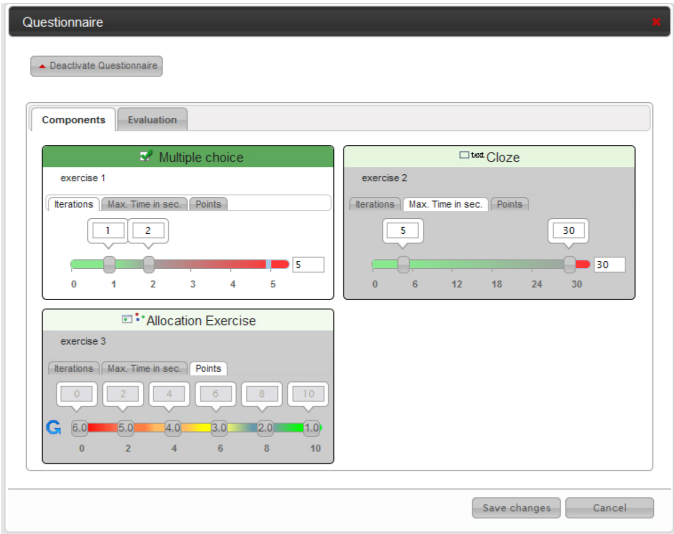

Figure 5. Example parameter setting for criteria certainty (tab “Iterations”), speed (tab “Max. Time in sec.”) and correctness (tab “Points”).

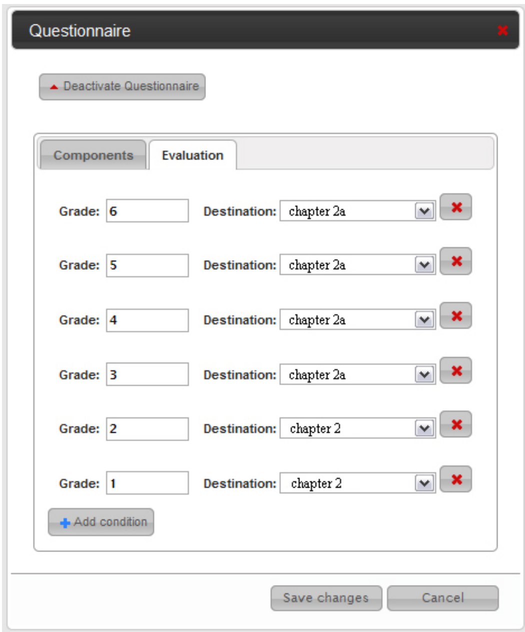

Figure 6. Linking grades to learning units.

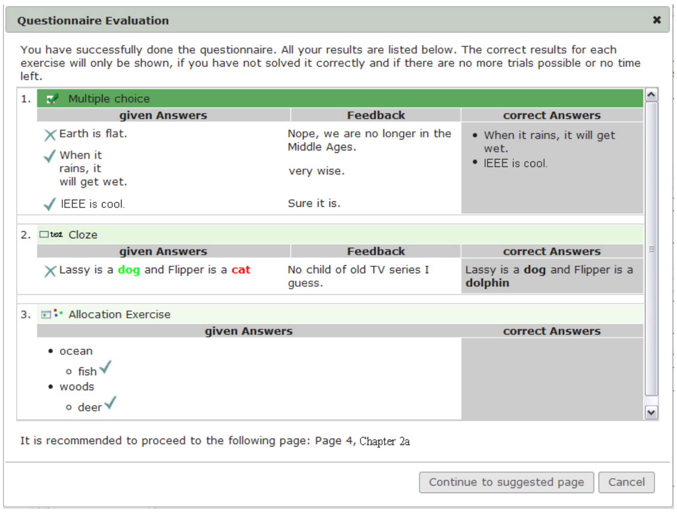

Figure 7. Test evaluation screen.

We are well aware of the discussion about whether it is useful to give a student more feedback than the so called knowledge of correctness or if additional feedback is reasonable [16]. However we believe that students want to be informed about their performance, so in each exercise students are presented with the maximum of time and iterations and how much time and iterations are left. We decided to pass on the decision about additional feedback to the author. The authoring front end therefore enables the author to write additional feedback – if he thinks it is useful. When the student finishes the test an evaluation window shows him the correct solution for each exercise and (if specified by the author) related feedback (see Fig. 7).

The implementation of the graphical user interface (GUI) uses AJAX technologies and jQuery library functions [17]. Data is communicated by JSON (Java Script Object Notation) objects [18] between the front end and the Java back end.

C. *Evaluation of acceptance*

The evaluation process of our concept is in an early state. We conducted a preliminary field study about the general acceptance of our implementation of the adaptive learning paths with student participants. Of primary interest has been to check if there are any fundamental errors in the concept itself or in our implementation before focusing in detail on the enhanced criteria.

With a partner organization we were able to create a scenario, where we could test the concept. An adaptive e-learning course layout was developed by the partner organization for one of their trainings in the context of adult education. This course is part of a blended learning concept. The students had to work through the e-learning course at home via browser before the phase of attendance started. At the end of the phase of attendance is a certification exam. The authors mainly designed shortcuts through the course. So when a student showed good results in the test, the e-learning application suggested the shortcut. Students could follow the suggestion (and take the shortcut without really knowing that it would be a shortcut) or just follow the “normal” path.

TABLE I. RESULTS VIA STANDARDIZED QUESTIONNAIRE (N=14)

question	yes	no	n.s.
Did you follow the suggested path?	11	3	0
Did you change your behavior in following (or not following) the suggested path during the course?	1	12	1
Did the suggestions motivate you for learning?	7	5	2
Do you think that your learning success with adaptive learning paths was better compared to a course with rigid paths?	4	9	1

The evaluation was conducted with n = 14 students. At the beginning of the phase of attendance they were asked several questions with a standardized questionnaire. Table I shows an excerpt of the results. At first glance it seems that adaptive learning paths have been accepted. A majority of the participants stated that they had followed the suggested paths and didn’t change that behavior. Some of the participants even thought that the adaptive learning paths increased their motivation for learning. However, when we look at the last question we see a slightly different picture. A majority stated that their learning success was not improved. When asked about the reasons of their judgment, some interesting statements were made. Some students didn’t seem to really trust the application, as the following statements indicate: “I was worried, that I could get pass something when taking the shortcut.” or “As a precaution I worked everything through.” or “Software cannot be as individual as a real trainer.” Some students obviously tricked the e-learning application in following the path but jumping back to the “shortened” chapters when they realized the shortcut.

A conclusion of this first evaluation is that the reasons for the system-made decisions have not been made sufficiently transparent to the students. So in a second attempt we will present more details about the reasons why the suggested path has been chosen. On the other hand it might be interesting what students think about learning success when they don’t know about the shortcuts and are just forced to follow the suggested paths. This scenario seems appropriate because a reason for working everything through might also be the fear of failing the certification exam.

As this group was rather small more evaluations with larger participant numbers combined with the control group methodology are planned. Further evaluations regarding more transparency of the decisions of the application as well as variations in the test scenario are also reasonable. Nevertheless this first evaluation clearly indicates a general acceptance of the web based adaptive learning paths.

D. *Visualization for the user*

Presenting alternative routes and adaptive learning paths for the user should also be evaluated in the future. Learning maps as specialized knowledge maps may be used and adapted according to the learning content for the visualization. The metaphor of a map as shown in Fig. 8 is useful for an adaptive e-learning system, as it assists its network-character [19]. And modern web technologies like the proposed SVG content integration in HTML 5 [20] may offer a wide range of possibilities in implementing a web based visualization of adaptive learning paths.

III. CONCLUSION

This paper presents a concept for an enhanced evaluation of tests to be used in web based adaptive learning paths. In addition to the obvious criterion correctness the additional criteria speed, certainty and difficulty level have been considered as valuable enhancements to determine the quality of a student's test result. The concept shows how to convert this enhancement technically for the decision process. The concept permits finding the adequate path through the learning course for each student in order to enhance the individual learning success. The technical foundation of this evaluation logic is the use of fuzzy logic elements as this allows the dedicated enabling and disabling of particular criteria. Furthermore it permits other criteria, especially "mathematically weak" criteria like for example the mood of a student, if this is considered a necessary criteria [21].

The presented prototype implemented the criteria correctness, speed and certainty according to the proposed concept. The prototype shows some suggestions how to design a web based GUI for authors dealing with the comparatively complex development of tests and structures for adaptive learning paths based on fuzzy logic. It also shows some ideas how to present the tests to the students and how to aggravate possible cheating.

A first evaluation of the concept and its implementation indicate a general acceptance of web based adaptive learning paths by students. The evaluation showed that such a concept does make sense and is worth to be evaluated in more detail. The next steps will focus on more transparency for students in terms of the decision made by the system. This is supposed to increase the acceptance rate of adaptive systems. The subsequent step is to focus more on the enhanced criteria, i.e. to compare groups using adaptive learning paths with enhanced criteria recognition and groups using adaptive learning paths with only the criterion correctness. A comparative study of the implemented prototype and an alternative adaptive e-learning system like ILIAS[3] may also be of interest.

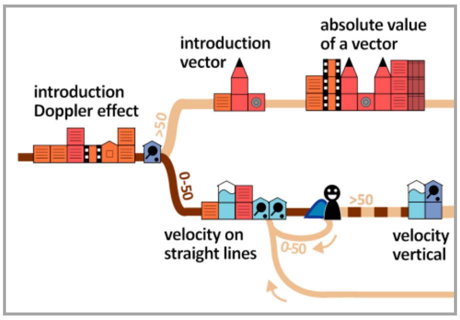

Figure 8. Example of a learning map showing alternative routes, different learning and knowledge units, walked path and actual learning 'position'.

[3] ILIAS® is an open source learning management system developed by the University of Cologne (https://www.ilias.uni-koeln.de).

REFERENCES

[1] W. F. Hill, "Learning: A survey of psychological interpretations," The Chaucer Press, Bungay, Suffolk, UK, 1977.

[2] P. Saettler, "The evolution of American educational technology," Information Age Publishing Inc., USA, 2004.

[3] W. Roller, B.-A. Bargel, P. Henning, F. Reinert and D. Szentes, "Methoden und Konzepte zur situationsangepassten Ausbildung und Assistenz," Proceedings of 4. Fernausbildungskongress der Bundeswehr, Ziel Verlag, Augsburg, Germany, in press.

[4] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz and M. von Wilamowitz-Moellendorff, „Gumo – the general user model ontology," Proceedings of the 10th International Conference on User Modeling, LNAI 3538: Springer, Berlin Heidelberg, pp. 428-432, 2005.

[5] F. Cretton and A. Le Calvé, "Generic ontology based user model: GenOUM," Centre Universitaire D'Informatique, Université de Genève, 2008.

[6] B. P. Woolf, "Building intelligent interactive tutors student-centered strategies for revolutionizing e-learning," Morgan Kaufmann Publishers, Burlington, USA, 2009.

[7] S. Baldiris, et al., "Integration of educational specifications and standards to support adaptive learning scenarios in ADAPTAPlan," in International Journal of Computer Science & Applications, vol. 5, pp. 88-107, 2008.

[8] M. Brut, S. Buraga, S. Dumitriu, G. Grigoras and M. Girdea, "A competency-oriented modeling approach for personalized e-learning systems," Proceedings of the 2008 Third International Conference on Internet and Web Applications and Services, IEEE Computer Society, pp. 410-415, 2008.

[9] A. Nitchot, L. Gilbert and G. B. Wills, "Towards a competence based system for recommending study materials (CBSR)," 10th IEEE International Conference on Advanced Learning Technologies, pp. 629-631, 2010.

[10] G. A. Lienert, U. Raatz, " Testaufbau und Testanalyse," Psychologie Verlags Union, Weinheim, Germany, 1998.

[11] J. Millman, C. H. Bishop and R. Ebel, "An analysis of Test-Wiseness", in Educational and Psychological Measurement, vol. 25, no. 3, pp. 707-726, 1965.

[12] W. Roller, B.-A. Bargel, "Intelligent cloze exercise," LearnTec 2010, talk in section 5 Semantic Technologies, 2010.

[13] A. Schreiber, "CBT-Anwendungen professionell entwickeln," Springer Verlag, Berlin Heidelberg, Germany, 1998.

[14] L. A. Zadeh, „Fuzzy Sets," in Information and Control, vol. 8, no. 3, pp. 338-353, 1965.

[15] D. Knaus, "Konzeption adaptiver Lernpfade in Web based Training Modulen und Beispielrealisation auf Basis des Content Management Systems WebGenesis," degree dissertation, University of Applied Sciences Karlsruhe, Germany, 2006.

[16] B. Jacobs, "Übungsaufgaben mit antwortabhängiger Rückmeldung," online article, URL: http://psydok.sulb.uni-saarland.de/volltexte/2005/516/html/adaptiv.html (last visit 17th of June 2011), 2003.

[17] J. Chaffer, K. Swedberg, „jQuery reference Guide," Packt Publishing Ltd., Birmingham, UK, 2007.

[18] D. Crockford, "The application/json Media Type for JavaScript Object Notation (JSON)," Request for Comments: 4627, July 2006.

[19] B.-A. Bargel, J. Schröck, D. Szentes, W. Roller, „Using learning maps for visualization of adaptive learning path components," unpublished.

[20] W3C®, "HTML5 - A vocabulary and associated APIs for HTML and XHTML," W3C Working Draft 25 May 2011, URL: http://www.w3.org/TR/html5/the-map-element.html#svg-0 (last visit 29th of June 2011), 2011.

[21] C. M. Hettena, B. L. Ballif, „Effects of mood on learning," in Journal of Educational Psychology, vol. 73, no. 4, pp. 505-508, 1981.