Adaptive test based on a “study map”

Ing. Dita Dlabolová, Department of Informatics, Faculty of Business and Economics, Mendel University in Brno
dita.dlabolova@mendelu.cz

Abstract
The paper describes a prototype implementation of an adaptive test, which is based on a hierarchical structure representing curriculum of the course. The structure is called study map. The described implementation is designed and tested on a course Informatics for Economists II. The application implementing the test algorithm is web-based, accessible to anybody and serves as a supportive tool for exam-preparation for the students of the above mentioned course.

Key Words
Adaptive test, computer-assisted test, study map, mind map, knowledge representation, test assessment, Informatics for Economists II, XML.
Introduction

To overcome drawbacks of the most common type of computer test for testing knowledge – computer assisted test with closed questions, a proposal of an adaptive test was created. This proposal is based on a pedagogical experiment performed in summer semester 2012 at Faculty of Business and Economics at Mendel University in Brno. The aim of the experiment was to compare results of a written exam (computer assisted test with closed question) and an oral exam (interview with an expert) and thus to reveal the main problems and drawbacks of the current computer testing to be able to design an alternative type of test. The course of the experiment and the way of collecting and processing the data is described in details in a paper (Dlabolová, 2012). The result of the experiment with a detail description of proposed solution is in a paper (Dlabolová, Rybička, 2013). Let me pin-out the most important results, because they are actually the base for this paper.

The main problems of the current computer test can be comprised as:

- the questions in the tests are usually accidently chosen from different areas of the course and their answering is showing only abandoned pieces of knowledge,
- hence the result of the test says nothing about the structure of student’s knowledge or about an actual understanding of the topic,
- it is rather a knowledge of relations between the terms that indicates the understanding of the topic,
- background knowledge was seriously missing in astonishing amount of test.

Proposed solution is to represent the curriculum of a course in a structure very similar to mind map, called “study map” (there are differences from the mind map, but they are not significant for the purpose of this paper, for more details see the above mentioned Dlabolová, Rybička, 2013). The study map is also supported by another structure – a concept map representing relationships between terms contained in the course. The study map represents a hierarchical structure of a course, each node represents one course topic. The most general topics are nearest to the center, the most precise ones in the leaves. It is an oriented graph, for the current use it is simplified (some edges are omitted) to be a tree. The proposal of the test algorithm is based on the structure of the study map. The test items are bound with the node – the topic, each test items has to belong to some topic. The course of the test follows the branches of the tree, which means that the order of the questions in the test follows some part

---

1 In the moment of submission [18-10-13], the proceeeding containing the mentioned paper is above to be published.
of the curriculum of the course. The test is adaptive with following principle: when the node is accomplished, a more specific node is examined (depth-first search principle), in the case of not knowing the topic in the node, a sibling node is visited (breadth-first search principle). For now, the choice of following ancestor or sibling is done randomly. It has to be said that this is not the final form of the adaptive algorithm, in the final form the choice of the following node will also depend on the supporting concept map.

The study map is divided into layers according the importance of the knowledge – first layer is basic knowledge (testing is over when the basic knowledge is missing), second layer is the core knowledge, there may be third layer, which is an expert (A-student) knowledge not necessary for passing the test. Number of layers is only a current proposal, the number could change depending on results of the first implementation.

Few branches of the study map are tested during the test (the particular number of branches should be also specified by experiments). The result is depicted as a colored map.

**Objectives and methodology**

Objective of the paper is to describe a first experimental implementation of the adaptive test based on the study map. The current proposal has some problems and ambiguities, the data collected by this prototype should serve for their removing. Possible technologies are described and discussed in this paper as well.

For the purpose of the prototype, few simplifications of the proposal are done:

- hierarchy of the study map is a tree, so it can be simply represented in XML format (this simplification doesn’t affect results to be obtained),
- the following node (both in the depth-first search and in the breadth-first search) is chosen randomly,
- there is only an internal division of the nodes of the study map into the layers, e.g. in order to collect higher amount of data the end of the test does not depend on layers.

The prototype will be web-based, freely accessible and offered to the students currently enrolled (winter semester 2013/2014) in the course Informatics for Economists II, so they can use it as a tool for exam preparation. Since there are around 600 students enrolled in the course each winter semester (UIS, 2013), sufficient amount of collected data is expected.

The purpose of the prototype is to clarify following problems and answer the questions concerning the current proposal of the test algorithm:
- Only one question from each topic is posed, according the answer, the knowledge of whole topic is assessed. This principle is probably too strict.
- No skipping of items is possible, which is forcing the students to guess the answer, which is an aspect which should be omitted by the adaptive test.
- What is the threshold of visited nodes in the breadth-first search in case of a sequence of wrong answers?
- When to decide about the end of the test in the case of a “loop” of wrong and right answers?
- What will be the students’ reaction on this unusual kind of test? (only one test item is displayed in one moment, the number of items in the test is not given)

**Results**

At first, a mind map of a curriculum of the course Informatics for Economists II was created. The map captures all significant topics taught during the course, which are important for the student to know. The map presented in this paper differs from the very first design, the topics are less disintegrated, which enables focus at the most important areas and using more questions from one topic is possible without causing the test to be unbearably long.

1 Study map of a course Informatics for Economists II
The study map representing the curriculum was created in open source software Freemind\(^2\), which stores maps in a MM format, which is XML. Freemind is an open mind-mapping software (under GNU/GPL), which is widely used. The source code of the map is very simple for basic maps (as the one at figure 1), hence for now it was an ideal solution. Unfortunately does not support optional attributes of nodes and creating of edges over the tree structure is complicated. For following use software called Freeplane\(^3\) (based on Freemind, GNU/GPL) will be used – it was designed to overcome the above-mentioned drawbacks of Freemind (Freeplane, 2013). Unfortunately source code of the map is less clear and for now the manipulation with the attributes does not work properly.

Since it is necessary to keep the relations also between the test items, the best solution is to comprise them into the nodes of the study map. This is very simple thanks to the XML representation: into particular nodes of the map are simply included nodes with the test items and the options. A structure of such XML follows, it is simplified – some nodes are omitted, hence the following code serves only for an illustration purpose of this paper:

```xml
<map>
  <node ID="1" TEXT="Central node">
    <node ID="2" TEXT="First ancestor">
      <item id="q1" text="First test item">
        <option id="2o1" text="First option" correct="yes"/>
        <option id="1o2" text="Second option" correct="no"/>
      </item>
      ...
    </node>
  </node>
</map>
```

Since this implementation will be used as an exam preparation for the students, old and currently not unused test questions will be used.

The test itself is implemented as a web application, php language and jQuery are used. The test is directly accessible without need to log-in. It can be accessed at http://akela.mendelu.cz/~dlabolo1/ipe2.

---

\(^2\) http://freemind.sourceforge.net
\(^3\) http://freeplane.sourceforge.net
Big problem of the first proposal of the algorithm was that it is too strict due to assessing the node only on a base of one test item. Since the first proposal of the study map was quite deep we did not want to use more items per node not to make the testing too long. We shrank the map, which enables us sing more items per node. Two test-items from each node will be used, if they are both wrong or both correct, next node follows, if their result is different, third item from the same node is posed.

When the test webpage is open and the test starts, only the first test item is depicted. The following item is depicted after the previous one is answered, hence only one item is displayed at one moment. It could be possible to display the two first items from one node together, but we do not want to confuse students with displaying different number of items (in the case of using the third “judge” item.

There are three different conditions under which the test is over:

- A leaf is reached.
- All ancestors of a correctly answered node are visited (i.e. all items from the ancestors were answered wrong).
- 20 test items were used (for the case of too long loop of wrong and right questions).

If six or less questions are used when the first two conditions are fulfilled, another branch of the map is tested – the reason is not to discourage student from next tries by too short test.

The current proposal of the algorithm does not enable skipping the test items, which can be limiting and students may be forced to guessing. This problem should be also examined by this prototype solution. We decided to provide an option “I don’t know, I don’t want to answer.”, but for the algorithm it acts like a wrong answer.

When the last test item is answered, a message informing about the end of the test is displayed and student is asked to evaluate himself. This is a principle similar to certainty-based marking (CBM)\(^4\). In CBM student should mark at each test item how sure he is with the answer, the amount of points he receives for the answer is based on this level of certainty. We were considering using the CBM principle (without the impact on the amount of points) during the test to get more feedback, but we decided to ask the student for the feedback just once. At first we hope students would pay more attention and would be more honest when they have to assess themselves only once, at second – the test as a whole is more important for us than particular test items. The choices are: “I’m an expert.”, “I know quite lot.”, “I knew a half of the test.”, “I didn’t pass, I don’t know the topic.”

\(^4\) Short description of CBM principle can be found e.g. here: [http://www.tmedwin.net/cbm/moodle/](http://www.tmedwin.net/cbm/moodle/)
Subsequently, the test is really over, all items (including chosen and correct answers) and an amount of points gained in the test are displayed. The reason for using a the scalar value as a test result (which is actually under critics in this paper) is simple – students are used to have a scalar value as a feedback of a test and we do not want to confuse them by presenting too many things at once. At this page with the result, we ask student for an optional verbal comment on the test.

Result of each test is stored into a database with a unique ID, which is the only identifier of the test. Other data stored on test are the test duration (time of the beginning and of the end), number of points, student’s self-evaluation and his comment.

Very important part for the future development is a non-scalar representation of the test result. The original mind map serves this purpose. The items are placed into the map at the nodes, where they belong. Correctness of the answer is marked by a color (green for a correct one, red for a wrong one, yellow when student didn’t know).

**Discussion**

The first overview on the test algorithm would be obtained from the comparison of the student’s self-assessment with the graphical representation of the result, also compared with the number of points obtained in the test. Subsequently the graphical representations of the test results will be examined.

Thanks to enabling to not answer the question, we would be able to assess the impact of skipped question and to decide if the implemented solution (take the skipped item as a wrong answered) is suitable.

Each test has a different length, an important step is comparing the results of tests with different lengths.

**Conclusion**

For future development of an adaptive test, which would be able to assess knowledge of a topic as whole, a simplified prototype solution was created. This adaptive test is based on so called “study map”, which is a hierarchical structure representing a curriculum of a course. Each node of the structure represents one topic of the course, each test item is bound to one particular node. The course of the test follows particular branches of the map. For the purpose of the prototype a course Informatics for economists II was used and the test is going to be used by students who are enrolled into this course. The designed solution contains a
background study map, XML database with stored test items, web-based application running the test and also a web-based depiction of the test results.

Unfortunately obtaining significant amount of data from students to be assessed is expected by the end of the semester at the beginning of the exam period.

The objective of this prototype is to show actual behavior of the adaptive test in the practice and provide knowledge for the future development, when the algorithm will be more specific, following more criteria for choosing following test item.

References


