# HIERARCHICAL SYSTEM FOR EVALUATING TESTS IN AN E-LEARNING SYSTEM USING AN EXPERT SYSTEM

**Bogdan Walek[1], Vladimir Bradac[2], Ondrej Pektor[2]**

[1]*Institute for Research and Applications of Fuzzy Modeling, University of Ostrava (CZECH REPUBLIC)*
[2]*Department of Informatics and Computers, University of Ostrava (CZECH REPUBLIC)*

## Abstract

The article deals with a hierarchical system and an approach for evaluating tests in e-learning systems using an expert system. Tests in e-learning system can be created in various modifications. In the proposed approach, the test in e-learning system is divided into categories of questions which represent the set of related questions. Then, we identify the criteria to evaluate questions and categories of questions. Then, we propose a hierarchical system to evaluate each question or a category of questions including continuous visualisation of the results. The hierarchical system uses an expert system for the evaluation. Based on visualised results, the creator of the test can recognise results for each question and the category of questions. The final evaluation of the e-learning test is also visualised and shown to the test creator. In the paper, individual steps of the proposed hierarchical systems are described. The proposed expert system is experimentally verified on test results of several students.

Keywords: e-learning, test, expert system, student, evaluation, hierarchical system, Takagi-Sugeno expert system.

## 1 INTRODUCTION

Nowadays, distance learning is very popular and offers useful methods of learning. Distance learning systems use e-learning methods of learning and teaching. E-learning is based on the availability of a system, which allows online courses to be developed and deployed [1]. It offers several advantages, especially concerning the management of the courses content [2], allowing the learning paths to be adapted to student's needs. There are also several intelligent systems for e-learning [3] [4] [5] [6].

E-learning systems usually contain a final course test or a final test of the entire content of learning courses or materials. Sometimes, e-learning systems also include ongoing (particular) tests in individual courses.

Tests in e-learning systems consist of questions and answers. For each question, there is one or more correct answers. Students gradually answer these questions. After answering all questions, the test is evaluated.

This paper proposes a hierarchical system for evaluating tests in e-learning system using an expert system. The proposed system is suitable for didactic tests and tests with questions divided into several categories. This article uses some principles published in our paper, which proposed a new approach to adaptive e-learning systems [7].

## 2 METHODOLOGY

The main goal of the proposed hierarchical system for evaluating tests in an e-learning system is to provide comprehensive information about test results. The creator of the test can recognise not only the results of test, but also results for specific categories of the test and results for all questions in a test category.

The structure of the proposed system is shown in Figure 1:

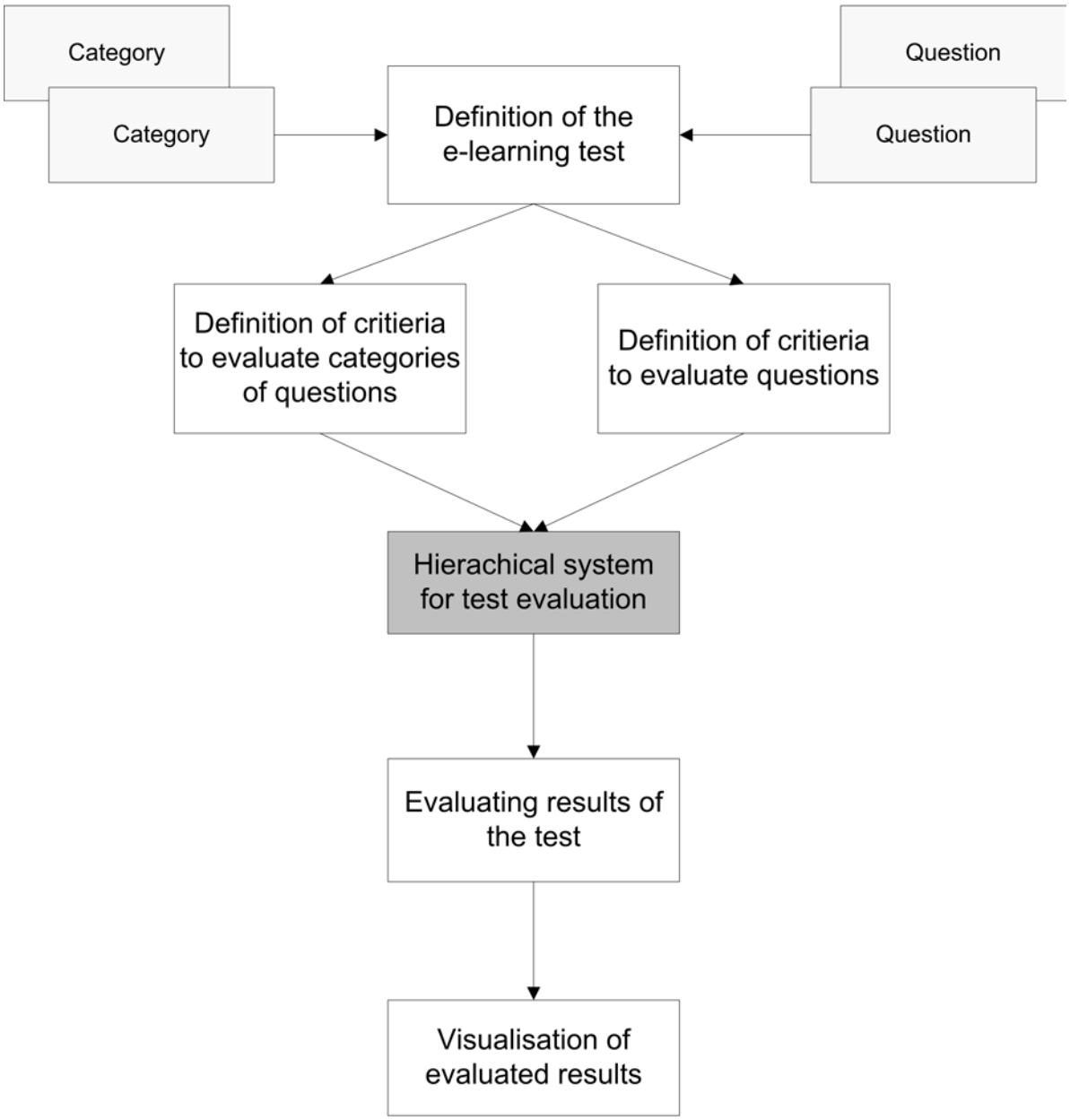


*Figure 1. Structure of the proposed system*

The main parts of the proposed system will be described in the following sections.

## 2.1 Definition of the e-learning test

In the first step, it is necessary to define and create an e-learning test. As we mentioned in the introduction, the division of the questions into categories is required. A category of questions often represents a set of related questions or questions with the same topic. Each category of questions consists of any number of questions.

Examples of several categories of questions in an e-learning system to learn English language are:

- Present perfect simple
- Past simple
- Present simple
- Present continuous
- Future tenses

- Conjunctions
- Numerals
- Adjectives, etc.

Having defined categories, the questions with possible answers and correct answer(s) are defined.

## 2.2 Definition of criteria to evaluate categories of questions

Next, the criteria for evaluating categories of questions are defined. In our system [7], each category of questions has an important attribute – importance of the category for further studies. This attribute represents how important the category is for further study (study in subsequent continuing courses). The attribute has a value from interval <0,10>; 0 – lowest importance, 10 – highest importance. Similarly, each question has an attribute called weight. This attribute represents how important the question is in the category of questions. It was created to distinguish more or less important questions in a specific category. The attribute has a value from interval <5,10>; 5 – lowest weight (default), 10 – highest height. During the completion of the test, we also measure how long it takes for a student to answer a specific question. This information is subsequently important to detect which questions are easy (student answers in a short time) or difficult (student thinks about answer for a long time) for a specific student. During the verification of the system, we set the time spent over an answer to 20 seconds. It means that an average student who knows the answer to a question spent 20 seconds at the most over the answer to a question. Based on these properties, the criteria for evaluating categories of questions were selected:

- Number of correct answers within the category
- Weight of correct answers within the category – average weight of correct answers
- Importance of the category for further studies
- Time spent over answers of the category – average time spent over answers of all question in a specific category

## 2.3 Definition of criteria to evaluate questions

Next, the criteria for evaluating questions are defined. The criteria for evaluating the questions are based on some of the properties defined in previous step. Important attributes of a question are: weight of question and time spent over answer of the question. And, of course, we need the information if the student´s answer for a specific question was correct or incorrect. Based on these attributes, the following criteria for evaluating questions were selected:

- Weight (importance) of the question
- Time spent over an answer to the question
- Correct or incorrect answer of the question

## 2.4 Creating a hierarchical system for test evaluation

In the next step, a hierarchical system for test evaluation is created. The hierarchical system consists of two parts:

- S1 – A system for evaluating categories of questions using a linguistic description with fuzzy IF-THEN rules
- S2 – A system for evaluating categories based on IF-THEN rules of a Takagi-Sugeno expert system

System S1 evaluates the overall evaluation of each category of questions. The output of system S1 is the evaluation for each category. Higher evaluation represents a higher success rate of student´s answers in a specific category based on the defined criteria for evaluation.

Input linguistic variables for S1 are:

- V1 - Number of correct answers within the category
- V2 - Weight of correct answers within the category – average weight of correct answers
- V3 - Importance of the category for further studies

- V4 - Time spent over answers of the category – average time spent over answers of all question in the specific category

Output linguistic variable is:

- V5 - Overall evaluation

Here are examples of IF-THEN rules of a linguistic description of system S1:

```
IF (V1 is small) and (V2 is small) and (V3 is very small) and (V4 is big)
THEN (V5 is extremely small)

IF (V1 is small) and (V2 is small) and (V3 is very small) and (V4 is small)
THEN (V5 is small)

IF (V1 is medium) and (V2 is small) and (V3 is medium) and (V4 is medium)
THEN (V5 is medium)

IF (V1 is medium) and (V2 is big) and (V3 is medium) and (V4 is big)
THEN (V5 is more or less medium)

IF (V1 is big) and (V2 is medium) and (V3 is big) and (V4 is medium)
THEN (V5 is very big)

IF (V1 is big) and (V2 is big) and (V3 is very big) and (V4 is small)
THEN (V5 is extremely big)
```

The creation of the linguistic description of the S1 system was performed in the LFL Controller. Linguistic Fuzzy Logic Controller is more described in [8].

System S2 evaluates the overall evaluation of each question. The output of system S2 is the evaluation for each question. The S2 system uses IF-THEN rules of the Takagi-Sugeno expert system for evaluation. Higher evaluation represents a higher success rate of student´s answers in specific question based on the defined criteria for evaluation.

Input linguistic variables for S2 are:

- WEIGHT - Weight (importance) of question
- TIME - Time spent over answer of the question
- ANSWER - Correct of incorrect answer of the question

Output linguistic variable is:

- EVALUATION - Overall evaluation

Here are examples of IF-THEN rules of linguistic description of S2:

```
IF (WEIGHT is low) and (TIME is low) and (ANSWER is wrong)
THEN (EVALUATION is 0.03)

IF (WEIGHT is medium) and (TIME is medium) and (ANSWER is wrong)
THEN (EVALUATION is 0.1)

IF (WEIGHT is low) and (TIME is low) and (ANSWER is right)
THEN (EVALUATION is 0.48)

IF (WEIGHT is medium) and (TIME is medium) and (ANSWER is right)
THEN (EVALUATION is 0.82)

IF (WEIGHT is high) and (TIME is low) and (ANSWER is right)
THEN (EVALUATION is 0.98)
```

## 2.5 Evaluating results of the test

In this step, the results for each student are evaluated. After completing the test, test results are stored in a database. Then, the results for all categories of questions and specific questions are inputs for system S1 and S2. The S1 system evaluates the overall evaluation of each category of questions in student´s test results. The S2 system evaluates the overall evaluation of each question in student´s test results.

## 2.6 Visualisation of evaluated results

Finally, the evaluated results are shown to the teacher (creator of the e-learning test). The evaluated results are visualised in a comprehensive form, so the teacher sees the evaluation of all categories of questions. The evaluation of all test questions and the results in specific categories or questions should be compared with each other. The teacher is also able to compare the test results and their specific parts (for instance, success rate of two or three selected questions) for all students in the e-learning course.

# 3 RESULTS

The proposed system has been experimentally verified on two groups of students who participated in an e-learning course KIP/ANGI3 at the Department of Informatics and Computers, University of Ostrava. The first tested group took the course in the winter semester 2014/2015, the second group in the winter semester 2015/2016. Both groups counted 16 students.

Figure 2 shows results of four categories and their questions of the first student.

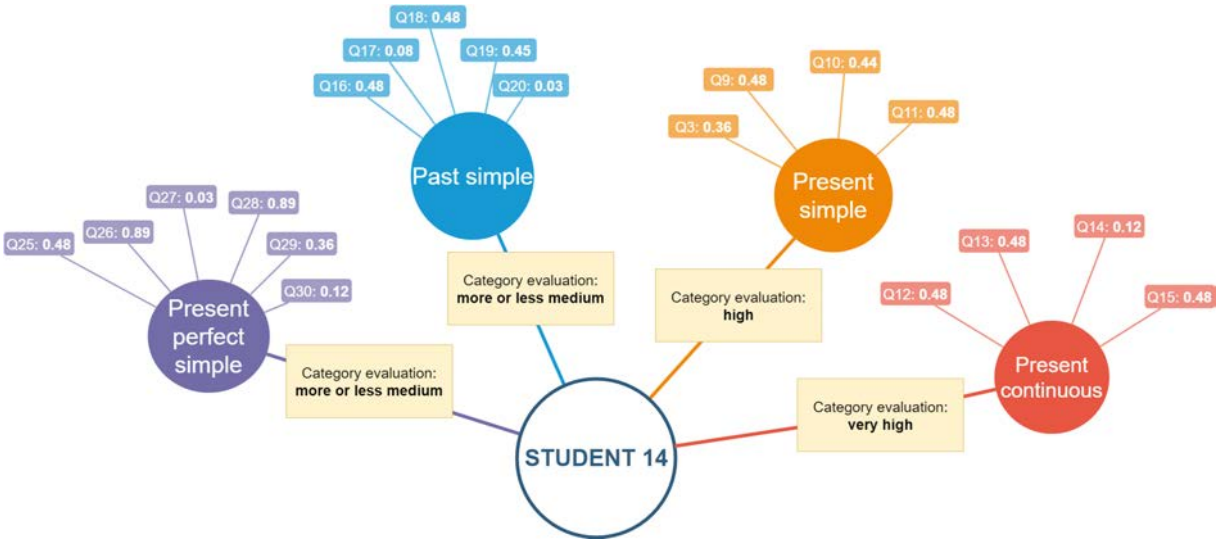


*Figure 2. Test results for the first student*

As we can see in Figure 2, there are four categories of questions and their overall evaluation. The teacher can thus see the overall evaluation of specific categories of questions for a student with ID value 14. The overall evaluation for each category is provided by the S1 system. For each category, there is also evaluation of all answers for a question in a specific category. We can see six results for category Present perfect simple, five results for category Past simple, four results for category Present simple and four results for category Present continuous. Each result represents overall evaluation of a question based on the weight of the question, time spent over the answer to the question and information if the answer was correct or incorrect. The result is a numeric value from interval <0,1>, higher value represents a better result for each question. Figure 2 provides a conclusion that results for questions Q26 or Q28 are very good, but results for question Q27 or Q20 are very weak.

Figures 3 and 4 represent results for next 2 students.

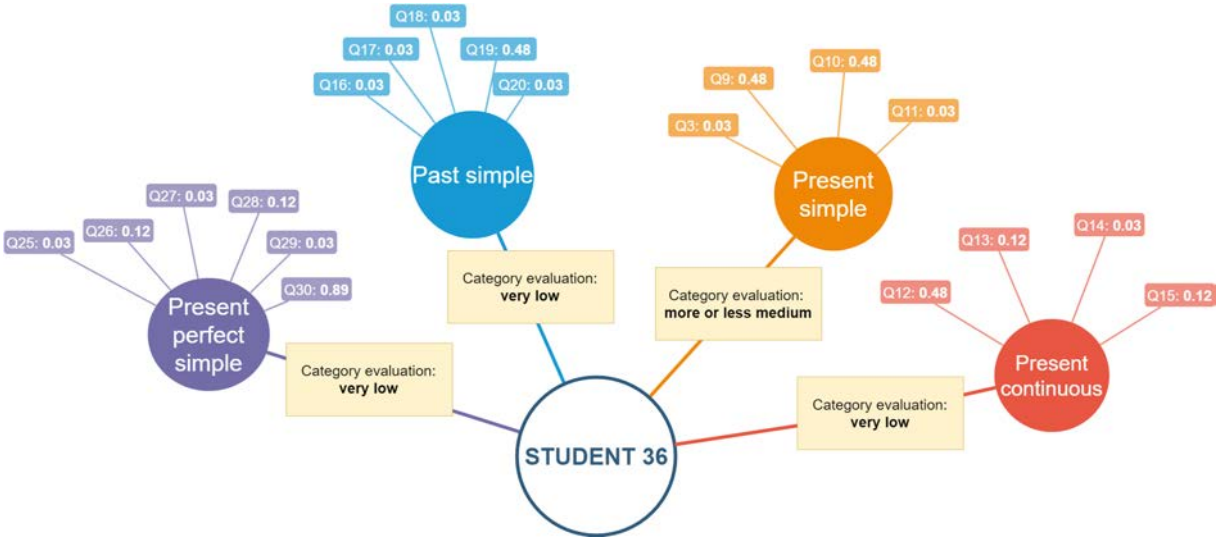


*Figure 3. Test results for the second student*

In Figure 3, we can see that the overall category evaluation of category Present perfect simple is very low because almost all answers in this category are incorrect. There is only one correct answer with high question evaluation – Q30. A similar situation is in category Present continuous. We can also see that three categories have evaluation “very low” so student´s results in this part of the course test are quite weak.

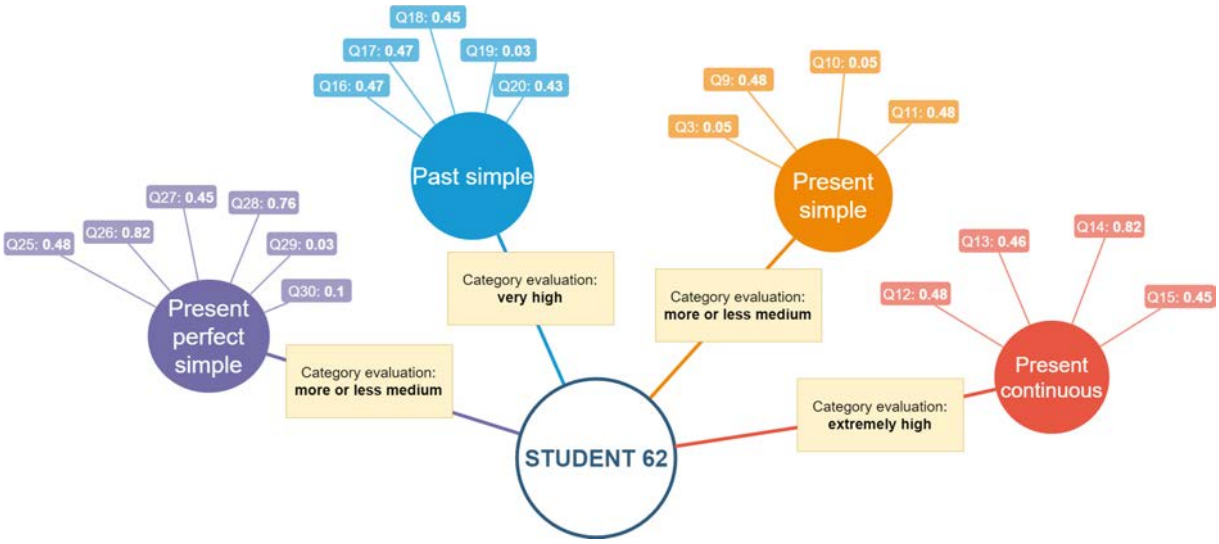


*Figure 4. Test results for the third student*

In Figure 4, there are results for the student ID 62. We can see that category Past simple has evaluation “very high”, because four questions were answered correct.

Based on the results shown above, the teacher (creator of the course test) is able to see comprehensive results of the e-learning test and to compare results of specific categories or questions between different students. The teacher also sees the information about the results of all questions, so he/she is able to conclude which questions are answered correct or incorrect in most cases. If the answers for a specific question are usually incorrect, it may imply that the question is so hard for students or even that the question was wrongly created.

## 4 CONCLUSIONS

The paper proposed a hierarchical system and an approach for evaluating tests in e-learning systems using an expert system. First, the e-learning tests and their evaluation where introduced. Then, the main parts of the proposed hierarchical system were described. The proposed system consists of two

parts: system S1 and system S2. The proposed system was experimentally verified on test results of several students.

Based on the visualised results of the course test, the teacher is able to see comprehensive results of the e-learning test and to compare results of specific categories or questions between different students. The teacher also sees information about the results of all questions, so he/she is able to conclude which questions are answered correct or incorrect in most cases.

## ACKNOWLEDGEMENTS

This work was partially supported by the NPU II project LQ1602 "IT4Innovations excellence in science" provided by the MŠMT. This work was supported by a Student Grant SGS01/UVAFM/2019 with student participation, supported by the Czech Ministry of Education, Youth and Sports.

## REFERENCES

[1] A. Nicola, M. Missikoff, F. Schiappelli, “Towards an Ontological Support for eLearning Courses”, On the Move to Meaningful Internet Systems 2004: OTM 2004 Workshops, Springer Berlin Heidelberg, pp. 773-777, 2004.

[2] T. R. Gruber, “A translation approach to portable ontologies”, Knowledge Acquisition, pp. 199-220, 1993.

[3] M. J. Huang, H. S. Huang, M. Y. Chen, “Constructing a personalized e-learning system based on genetic algorithm and case-based reasoning approach”, Expert Systems with Applications, vol. 33, no. 3, pp. 551-564, 2007.

[4] P. Tzouveli, P. Mylonas, S. Kollias, “An intelligent e-learning system based on learner profiling and learning resources adaptation”, Computers & Education, vol. 51, no. 1, pp. 224-238, 2008.

[5] B. P. Woolf, “Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning”, Burlington, Elsevier Inc., 2009.

[6] A. Gladun, J. Rogushina, F. Garcia-Sanchez, R. Martínez-Béjar, J. T. Fernández-Breis, “An application of intelligent techniques and semantic web technologies in e-learning environments”, Expert Systems with Applications, vol. 36, no. 2, pp. 1922-1931, 2009.

[7] V. Bradac, B. Walek. "A comprehensive adaptive system for e-learning of foreign languages" Expert Systems with Applications, Vol. 90, pp. 414-426. 2017.

[8] H. Habiballa, V. Novák, A. Dvořák, V. Pavliska, “Using software package LFLC 2000”, 2nd International Conference Aplimat 2003, Bratislava, pp. 355-358, 2003.