Testing of Understanding by Use of an Ontology Methodology

Larisa Soldatova, Riichiro Mizoguchi
ISIR, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan
{larisa, miz}@ei.sanken.osaka-u.ac.jp

Abstract: *We describe a test composition approach to determine the level of a student's understanding, based on a test ontology. Domain independent and domain dependent parts of the student's model are described. To improve the diagnostic of student knowledge we used the 6-layers Bloom's model of understanding. Reusable components and steps of the test construction are listed. This approach is an example of "function integration". To validate the approach a pedagogical experiment was carried out at Vladivostok State University, Russia. The results are discussed. The proposed method can be used for providing a regular feedback in intelligent educational systems.*

Key words: *ontology, test/ problem generation, understanding, student model.*

Introduction

To successfully apply any IES (intelligent educational system) it is necessary to obtain feedback from a student about his/her level of understanding. Tests are the most popular method to measure a learner's knowledge. Unfortunately, standard types of tests and test scoring schemas do not generally provide enough information about a student understanding of study material (Crocker, L., & Algina, J., 1986; Steven, J. O., 1998).

The purpose of our research is the analysis of cognition process based an ontology methodology. Tests of understanding are a crucial part of any educational technology: traditional or based on information technology. We would like to automate the test composition and provide a diagnosis of the level of a student's understanding. For the purpose we use the popular pedagogy Bloom's 6-layers model of understanding. Each element of student knowledge model corresponds to the definite level of understanding. Using of the information about knowledge elements and relationships between them for test item generation allows us to perform a better diagnosis of student knowledge.

"An ontology provides us with effective methodology and vocabulary for both analyzing and synthesizing knowledge-based systems" (Marcke, V. K., 1995). An ontology can provide:

- a formalization of the description of test tasks and student model;
- a specification of the main concepts and relationships between them for the test construction process;
- reusable components for test design system;
- a shared vocabulary;
- and aid in the analysis of methods of test composition.

Student Knowledge Base

We consider two parts of the knowledge base: domain independent and domain-dependent knowledge (Mizoguchi, R., & Sinitsa, K., 1996; Sowa, J.F., 2000). The domain-independent part consists of cognitive ability as reasoning, capability, recognition; general knowledge as methodological, structural, classification knowledge; and mental skills (see Fig.1, made in OE (Kozaki, K., Kitamura, Y., Ikeda, M., & Mizoguchi, R., 2000), where p/o is "part of" link with a slot, a/o is "attribute of" link, upper part of a slot corresponds to a role of the concept, and right side of a slot represents a class).

The domain-dependant knowledge part consists of nodes, such as facts, rules, theorems, terms, and principles; and links of different types (see Fig. 2).

Level of Understanding

In test composition we would like to use popular in pedagogy 6-layers model of understanding, suggested by Bloom (Bloom B.S. et. al., 1956). Compared to a standard "flat" student knowledge model, the 6-layers model provides more information about the student's stage and area of his/her difficulties. Each test item should correspond to the definite level of understanding and check the determined knowledge elements from the student model:

1st level, knowledge in a narrow sense, knowing facts and definitions of notions;

2nd level, comprehension, is understanding the meaning of notions, objects, abstracts, and knowing simple rules;

3rd level, application, corresponds to the ability to apply known rules;

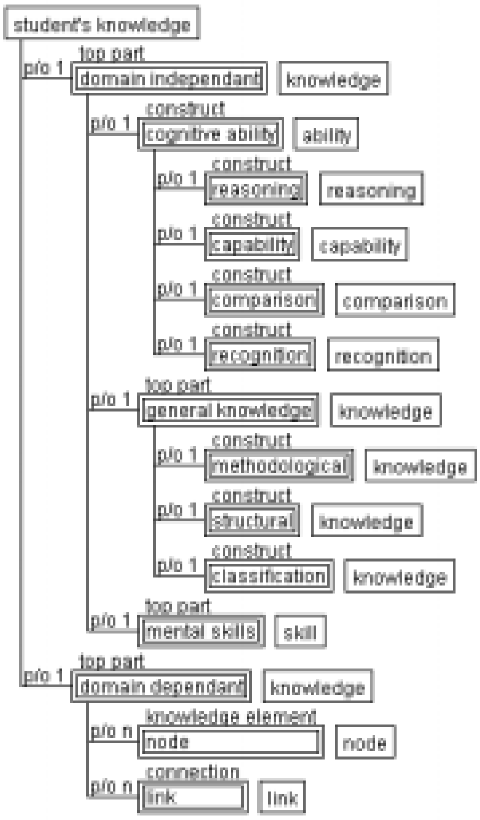


Fig. 1: Ontology of domain - independent knowledge (fragment).

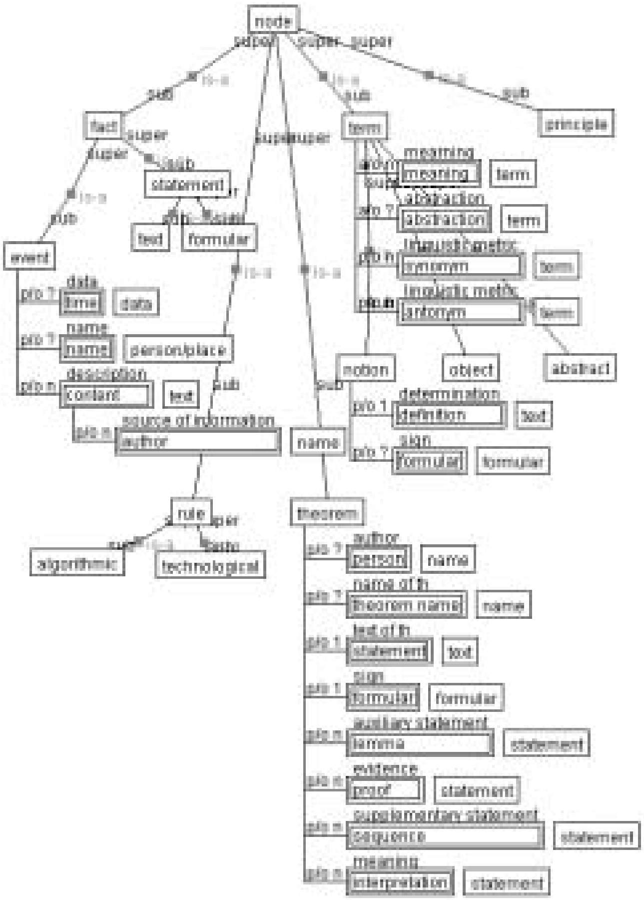


Fig. 2: Ontology of domain knowledge (fragment).

4th level, analysis, checks the understanding of relationships between elements, ability to select and compound different rules;

5th level, synthesis, corresponds to the ability to generalize knowledge;

and 6th level, evaluation, uses metaknowledge.

The first levels correspond to the domain knowledge, and upper levels reflect mostly the domain independent knowledge.

It is possible to give some recommendations about item forms for checking different layers of understanding (see table 1). We classify types of item forms as the following (Soldatova, L., & Mizoguchi, R., 2003):

1. Open question: fill-in-the-blank, give a short answer, and write essay.
2. The most popular form is multiple choice problems. We divide them according to amount of given answers (2, 3,…, etc.) and logic of answer (select one correct answer, select the most correct answer, or note all correct answers). The best known examples of such item form are true-false (T/F) questions (is the statement true or false) and SAN – questions (select: the statement is correct sometimes, always, or never).
3. Matching problems with different amount of columns.

For instance, fill-in-the-blank questions are good to check knowing facts, understanding of notion definition and meaning.

*Table 1: Recommendation for item forms, where sign **--** means not recommendable, * recommendable, ** very recommendable, and *** greatly recommendable.*

Level of understanding	Open question			Multiple choice			Matching
	Fill-in-the-blank	Short answer	Essay	T/F	SAN	Other	
I	***	**	--	**	*	*	*
II	***	***	--	***	**	**	***
III	*	***	--	**	**	***	***
IV	--	**	*	*	*	**	*
V	--	**	***	--	--	--	--
VI	--	*	***	--	--	--	--

From classical test theory we can extract several practical rules to determine for which levels of understanding and how many test items to construct for providing the best test properties (Anderson, L.W., & Krathwohl , D.R., 2001; Avanesov, V.S., 1999; Crocker ,L., & Algina, J., 1986).

Test Item Difficulty

In pedagogy one of the common approaches to measuring a test's item difficulty is the determination by experiment of the ratio of the amount of questions students answered correctly on an item divided by the total amount of the students (Avanesov, V.S., 1999). In (Soldatova, L., 1999) is discussed how to calculate the difficulty on the base of amount of used operations for the solution. However, sometimes a complicated problem does not require many different operations, but instead a deep understanding of the problem domain and advanced mental abilities.

The difficulty of a test item is related to the level of understanding required to answer it. It is therefore possible to carry out a classification of test items in a pool based on an analysis of the knowledge elements and mental skills used; or especially generate test items for the required level of understanding to check the required knowledge elements and relationships between them.

Approach for Test Construction

The method is based on the following reusable components (Soldatova, L., & Mizoguchi, R., 2003):

1. A test ontology with description of test tasks, purposes, types and kinds, scoring schemas, test items forms and so on.
2. A knowledge base which contains rules for test construction process.
3. A student model, domain - independent and domain – dependent parts.

The steps for the test construction are:

1. Detailing of the problem – dependant part of the student model.
2. Classification of test item pool according to the levels of understanding.
3. Determination of the test structure using the test ontology and the knowledge base, and generation or using ready test items according to the structure.

The result is a test which allows checking a level of student understanding of studying material.

Pedagogical Experiment

To check the suggested approach a pedagogical experiment on students of "management" and "accounting" specialties of Vladivostok State University of Economy and Service, Russia, was carried out. Function integration in the course "Calculus" as example of a problem domain was selected. The purpose of the test is to rank the students. The instructor's preferences are restriction of the test length by 10 test items and the test time by 40 minutes. We analyzed the problem domain, correspondence knowledge elements of the studying material to the levels of understanding and constructed a test by the methodology described in (Soldatova L., Mizoguchi R., 2003). On the basis of Test Ontology and a knowledge base about the test design process we produced the following test structure to provide good quality of the test (see table 2):

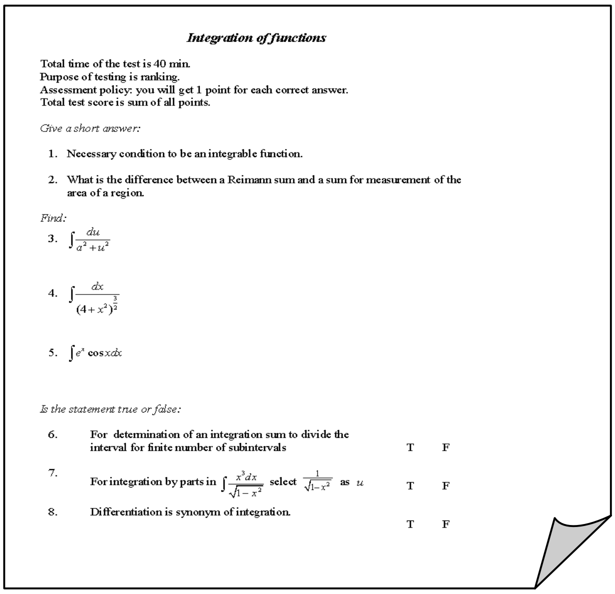
Table 2: Generated test structure.

Level of understanding	Open question			Multiple choice			Matching
	Fill-in-the-blank	Short answer	Essay	T/F	SAN	Other	
I				1		1	
II		1		2		1	
III		1		1		1	
IV		1					
V							
VI							
Total:		3		4		3	

The instructor of the students also created a test for the same purpose to compare the tests results. The structure of the instructor-made test did not allow us to check well students understanding; for example it contains 6 from 10 items of 3rd level of understanding from the suggested approach point of view. Moreover, the test included only trivial theoretical questions, which require only reproduction of known facts. Thus, a student who understands the study material, but is not good at calculation, may obtain a low test score and a place in the ranking.

You can see a fragment of the generated test by the suggested approach in figure 3. For the test composition were used the recommendable textbooks (Leithold, L., 1976; Piskunov, N.S., 2002).

It is possible to confirm that the test items generated by the suggested methodology correctly correspond to the level understanding because 90% of the students' profiles are correct (the students answered on simple questions and did not on difficult ones). Analyses of the test results have allowed us to precisely detect a lack of the 3rd level of understanding in the ability to apply known rules and at the 4th level in the ability to select and compound different rules. If we compare with the results of the instructor made test we can determine problems in students' understanding only intuitively because it is not clear which knowledge elements and of which level of understanding the test items check. Another positive effect of our test is the strong distribution of the test results near with the top of the ranking list, this is important for a placement test.



Integration of functions

Total time of the test is 40 min.
Purpose of testing is ranking.
Assessment policy: you will get 1 point for each correct answer.
Total test score is sum of all points.

Give a short answer:

1. Necessary condition to be an integrable function.
2. What is the difference between a Reimann sum and a sum for measurement of the area of a region.

Find:

3. $\int \frac{du}{a^2+u^2}$
4. $\int \frac{dx}{(4+x^2)^{\frac{3}{2}}}$
5. $\int e^x \cos x dx$

Is the statement true or false:

6. For determination of an integration sum to divide the interval for finite number of subintervals T F
7. For integration by parts in $\int \frac{x^3 dx}{\sqrt{1-x^2}}$ select $\frac{1}{\sqrt{1-x^2}}$ as u T F
8. Differentiation is synonym of integration. T F

Fig. 3: Test example (fragment).

We have not obtained enough statistics of pedagogical experiments to make a final conclusion, but we consider that the test generated by the our methodology better determines the level of students understanding than the traditional one, and additionally provides more diagnostic information, and takes less time for construction as it is partially automated.

Conclusion

In this paper a methodology for test construction to determine a student level of understanding is described. The methodology is based on a test ontology, a knowledge base about test design process, and the student model. The 6-layer Bloom's model of understanding helps to determine limitations in student understanding. The results of a pedagogical experiment provide evidence for the advantages of our approach compared with traditional methods. The methodology can also be used for providing feedback in IES or in traditional educational technology.

References

Anderson, L.W., & Krathwohl, D.R. (2001). *A Taxonomy for Learning, Teaching, and Assessing.* Addison Wesley Longman, Inc.

Avanesov, V.S. (1999). *Modern Methods of Learning and Control of Knowledge*. DVGTU, Vladivostok (in Russian).

Bloom, B.S., et. al. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Handbook 1: Cognitive Domain. N-Y, David McKay Co.

Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Wadsworth Group, Thomson Learning.

Kozaki, K., Kitamura, Y., Ikeda, M., & Mizoguchi, R. (2000). Development of an Environment for Building Ontologies which is based on a Fundamental Consideration of "Relationship" and "Role". *A workshop, Proc. PKAW*, Sydney, Australia, pp. 205-221.

Leithold, L. (1976). *The Calculus with analytic geometry*. Harper & Row, Publishers.

Marcke, V. K. (1995). Learner adaptivity in generic instructional strategies. *Proc. AIED*, pp. 323-333.

Mizoguchi, R., & Sinitsa, K. (1996). Task Ontology Design for Intelligent Educational/Training Systems. *Workshop on "Architectures and Methods for Designing Cost-Effective and Reusable ITSs", Proc. ITS'96*, Montreal, pp. 1-21.

Piskunov, N.S. (2002). *Differential and International Calculus*.V.1, Moscow, Integral-Press (in Russian).

Soldatova, L. (1999). Comparison of the Problems of Complicity. *Proc. ICCE, IOS Press, V. 2*, pp. 169-171.

Soldatova, L., & Mizoguchi, R. (2003). Test Generation Systems, *Proc. SIG-IES-A203-09* (Japan), pp. 51-56.

Sowa, J.F. (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks/Cole.

Steven, J. O. (1998). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats*. 2nd ed. Kluwer Academic Publishers.