The Educational Uses of Mathematical Ontology and the Searching Tool

Noriko Kitani, Shuichi Yukita
Hosei University, ZWQ03640@nifty.com, yukita@hosei.ac.jp

Abstract - In Japan, a huge number of entrance examination questions and school textbook exercises are accumulated in databases at various educational institutions and companies in education industry. These databases are now becoming more and more indispensable educational resources for teachers, educational researchers, and learners. However, since these resources are stored in a wide variety of forms, we can not search useful data at the semantic level. To tackle the problem, we constructed mathematical ontology and developed the problem searching tool based on this ontology in our previous work. We use standard ontology technologies which were developed for shared use and reuse of knowledge. The ontology captures one or more experts’ conceptual representation of a domain expressed in terms of concepts and the relationships among the concepts. We analyzed that the features of this technology could be the answer to solve above mentioned problem. In this paper, we describe an overview of the mathematical ontology and the searching tool, focusing on newly added functions, and demonstrate effective utilization of this tool in the field of mathematics education.

Index Terms – MathML, Ontology, Searching tool, Semantic Web, XML

INTRODUCTION

In Japan, a huge number of entrance examination questions and school textbook exercises are accumulated in databases at various educational institutions and companies in education industry, which are to serve teachers, students, and courseware designers in reusing educationally meaningful quality questions. These databases are now indispensable educational resources. Due to this environmental change, it is becoming more and more important to share the storage of questions electronically among distributed databases.

This research focuses on the sharing of the resources in mathematics. The big problem is that the questions are stored in incompatible forms and can not be searched at the semantic level. To effectively utilize the stored questions in classrooms, it is important for educators and students to retrieve questions by specifying learning goals, mathematical notions, the ways of giving questions, etc. To tackle the problem, we can base our work on the semantic web technology. Our first step is to build an ontology designed for our own purposes.

In our previous work, we described relationships between study items with references to school textbooks and Education Ministry guidelines using the ontology technology. We collected typical example problems related to study items from school textbooks and stored them in MathML. Integrating all these data, we constructed a mathematical ontology. Further, on top of this mathematical ontology, we developed a searching tool OMQS (Ontology based Mathematics Question Searcher) that supports searching by concepts or meanings independent of terminologies and presentations [1].

When the teacher wants to use past entrance examinations in class or reuse them in a term examination, searching by the difficulty level or complexity of questions is quite useful. We added this new feature to our system.

In this paper, we describe an overview of OMQS, focusing on newly added functions, and demonstrate effective utilization of this tool in the field of mathematics education. The second section gives an overview of the constructed mathematical ontology. The third section gives a brief introduction to the searching tool. The fourth section gives an overview of new features that include the capability to determine the difficulty level of computational problems. The fifth section gives an outline of a simple question input supporting tool. The sixth section gives a short description of educational environment and our approach. Concluding remarks are given in the final section.

MATHEMATICAL ONTOLOGY

The concept of ontology was introduced in the AI field to support the sharing and reuse of formally represented knowledge among different AI systems. An ontology describes the concepts and relationships that are important in a particular domain, providing a vocabulary for that domain as well as a computerized specification of the meaning of terms used in the vocabulary [2]-[4].

The ontology technology is best suited for our purpose: sharing information and reusing useful educational resources in mathematics.

We construct a mathematical ontology covering lower high school mathematics, which has a two-tired structure accommodating two levels of abstraction: the study item structural level and the study content level.

Figure 1 shows an overview of mathematical ontology.
The ontology shows the relationship among study items at the study item structural level and the included questions related to one study item at the study content level.

Study items are relatively stable because school textbooks are updated with the revision of Education Ministry guidelines, which takes place once in several years. On the other hand, materials at the study content level are rapidly accumulating because many questions including entrance examination are created every year. For this reason, we separated the ontology into two tiers: one tier stable and another subject to change [5]-[9].

Covering the whole resources using only one kind of ontology is difficult because there are many kinds of school textbooks and study books. So, at the study item structural level, we construct a small ontology corresponding to one school textbook or study book and integrate them into the entire ontology.

We represent a study item as a class in the ontology and relate it to equivalent items in the other ontologies according to Education Ministry guidelines and study item mapping tables. Using equivalence of classes, we can establish the fact that one study item in a school textbook is equivalent to another study item in another publication [10]-[15]. Interrelations are embedded via tag constructs.

The editing tool is necessary to construct Ontologies efficiently. We use Protégé for this purpose. Protégé is a widely used ontology editor developed at Stanford University, which is free, open source and supports multi-language [16].

For compatibility, we denote the contents of questions by MathML at the study content level. It also contains information about problem statement and other related information if necessary. We next associate these questions with study items of school textbooks or study books by tagging them appropriately.

MathML is a markup language for describing both the visual structures and the meanings of mathematical formulas. MathML supports two kinds of markup: Presentation markup and Content markup. Content markup describes the exact meaning of the mathematical expression [17]-[18].

We next show an example of content markup by MathML. An apply element typically applies an operator to its arguments at Content markup. The opening and closing tags of apply specify exactly the scope of any operator or function. A number is represented by <cn> tag and a character by <ci> tag [17]-[18].

Figure 2 shows an example of a simple polynomial expression \(x + (x + 1) + (x + 2) = 42\) with content markup.
We developed a question searching tool based on the mathematical ontology to prove the effectiveness of our semantic approach toward educational applications. We call the tool OMQS (Ontology based Mathematics Question Searcher).

Figure 5 shows the use case diagram of the question searching tool.

Question retrieval patterns are classified into three major types: Systematic Retrieval, Content-addressed retrieval and More complex Retrieval.

Table 1 shows the retrieval patterns with which teachers and teaching material authors usually search appropriate questions from storage. Items marked with a circle are currently supported. Those marked with an asterisk and those with a cross mark are partly supported and to be supported, respectively. For example, searching questions which involve a quadratic equation that has two real solutions lying between 0 and 1 is possible.

While conventional searching tools do not support searching by content of a mathematical expression or finding a quadratic expression with a specified set of answers, OMQS accommodates with those retrieval demands. OMQS displays the corresponding mathematical expressions by inputting search criteria, for example degree of expressions, algebraic signs and educational area.

DETERMINING THE DIFFICULTY LEVEL OF QUESTIONS

In a classroom, students usually have various academic backgrounds and teachers have to provide them with materials and achievement tests of appropriate levels. To cope with this situation, the retrieval by the difficulty level of questions is important for teachers and teaching material authors.

We add a new function, searching by the difficulty level of questions, into the searching tool to meet the above mentioned need. Especially, the new function supports the retrieval patterns by the difficulty level of mathematical expressions determined by the types and degrees of mathematical expressions and by the prerequisite knowledge. We show the details of this new function.

We first have to define the difficulty level of mathematical expressions from the view point of cognitive psychology and other educational researches. Apart from the systematic approach to difficulty level, there are many intuitively acceptable criteria for the difficulty level of mathematical expressions. For example, a division operation in which the dividend is larger than the divisor is considered more difficult than a division with a small divisor and a large dividend.

Many papers concerning a systematic approach are found in the literature. For example, the classification of
In the mathematical expression varies according to the location of unknown quantity.

"Change" questions, which involve a process whereby there is an event that alters the value of the quantity, for example: "Peter had 3 oranges. Michelle gave him 2 more oranges; how many oranges does Peter have now?"

"Combine" questions, which relate to static situations in which there are two amounts. These are considered either as separate entities or in relation to each other, for example: "Sarah has 4 oranges; Michelle has 2 oranges. How many oranges do they have altogether?"

"Compare" questions, which involve the comparison of two amounts and the difference between them, for example: "Ben has 5 oranges. Alice has 2 more oranges than Ben; how many oranges does Alice have?"

In addition to the above mentioned differences of the semantic structure, the difficulty level of word problems varies according to the location of unknown quantity.

In the mathematical expression “a + b = c” or “a − b = c”, the question is considered difficult in the following order: in the case where the unknown is quantity a, in the case where the unknown is b, and in the case where the unknown is c [19]-[20].

The difficulty level of questions is also affected by many other elements, for example, educational systems and educational policies of each school and teacher. So, we attach a supplementary parameter of difficulty level ranging from 1 to 5 to questions. By adding the parameter that controls the values of difficulty level, we can incorporate the retrieval pattern by the difficulty level determined by academic backgrounds and achievements.

We show a part of rules of determining the difficulty level of mathematical expressions.

Positive number and negative number

- The question becomes difficult as the number of operators increases. (The number of <apply> tag equals to the number of operator)
- Judgment by the kind of the operator. The order of the operator becomes difficult in the following order, + < − < × < ÷. Judge by the content of the tag following <apply> tag.
- The questions with decimal numbers or fractional numbers as a calculation result are more difficult than ones with only integer numbers.
  - Subtraction: The question with negative number as an answer is more difficult than one with positive number. Judge by calculating operators and numbers.
  - Division: Dividing calculation where the remainder has answers other than 0 is more difficult than one with 0 as a remainder. Judge by calculating operators and numbers.
  - Division: The question whose divide number is larger than the divided number is more difficult than one with inverse order condition. Judge by calculating operators and numbers.
- The more the number of digits increases, the more difficult the question is. Judge by the content of <mn> tag.

Character and Expression

- The question becomes difficult as the number of operators increases. Same rule as previously described.
- Judgment by the kind of the operator. The order of the operator becomes difficult in the following order, + < − < × < ÷. Same rule as previously described.
- The questions with decimal numbers or fractional numbers as constant terms and coefficients are more difficult than ones with only integer numbers. Judge by calculating operators and numbers.
- The questions with decimal numbers or fractional numbers as a calculation result are more difficult than ones with only integer numbers. Same rule as previously described.
- The larger the degree is, the more difficult the question is.
- The more variables and constant symbols are involved, the more difficult the question is. This function is not mounted at this time.
- The word problem is typically more difficult than a computational problem. This function is not mounted at this time.

Calculation of expression

- The question becomes difficult as the number of operators increases. Same rule as previously described.
- Judgment by the kind of the operator. The order of the operator becomes difficult in the following order, + < − < × < ÷. Same rule as previously described.
- The questions with decimal numbers or fractional numbers as constant terms and coefficients are more difficult than ones with only integer numbers. Same rule as previously described.
- The questions with decimal numbers or fractional numbers as a calculation result are more difficult than ones with only integer numbers. Same rule as previously described.
- The larger the degree is, the more difficult the question is. Same rule as previously described.
- The more variables and constant symbols are involved, the more difficult the question is. This function is not mounted at this time.

Polynomial Equation

- The question becomes difficult as the number of operators increases. Same rule as previously described.
- Judgment by the kind of the operator. The order of the operator becomes difficult in the following order, + < − < × < ÷. Same rule as previously described.
- The questions with decimal numbers or fractional numbers as constant terms and coefficients are more difficult than ones with only integer numbers. Same rule
The questions with decimal numbers or fractional numbers as a calculation result are more difficult than ones with only integer numbers. Same rule as previously described.

- The larger the degree is, the more difficult the question is. Same rule as previously described.
- The more variables and constant symbols are involved, the more difficult the question is. This function is not mounted at this time.
- The word problem is typically more difficult than a computational problem. This function is not mounted at this time.
- The calculations with nested structure are more difficult than ones with no nested structure. Judge whether there is \(<apply>\) tag in the scope of \(<apply>\) tag. (Judge by the number of \(<apply>\) tag.)
- Factorization: The expression that is hard to discover the common terms is more difficult than one otherwise. Judge by comparing the contents of the scope of \(<apply>\) tag.
- Judgment by kind of operator. The order of the operator becomes difficult in the following order, + < - < × < ÷. Same rule as previously described.
- The questions with decimal numbers or fractional numbers as constant terms and coefficients are more difficult than ones with only integer numbers. Judge by calculating operators and numbers.
- The word problem is typically more difficult than a computational problem. This function is not mounted at this time.

Figure 6 shows a snapshot of OMQS new version by adding the above mentioned new function.

We can search questions by specifying the difficulty out of the five levels. The system displays the questions that have the specified level of difficulty.

**SUPPROTING TOOL**

A large number of exercises and entrance examination questions are created at various educational institutions and schools every year. To incorporate the state of affairs, it is convenient for classroom teachers and learners to store new questions to the database. Although, there are many kinds of MathML editing tool, which must first be installed in the working environment and its instruction manual must be carefully read through. This requires some technical knowledge of software tools in general.

To cope with these problems, we provide a simple question input supporting tool. This tool helps the user input MathML portions of questions and answers. Figure 7 shows a snapshot of MathML input supporting tool.

![MathML input supporting tool](image)

**FIGURE 7**

A SNAPSHOIT OF MATHML INPUT SUPPORTING TOOL

By using this tool, we can get the corresponding MathML notation and add the information into the existing mathematical ontology.

**EDUCATIONAL ENVIRONMENT AND OUR APPROACH**

As society changes and the educational environment becomes increasingly complex, conventional classification of mathematical contents into algebra, analysis, geometry, etc. does not meet our demand any more. Classification by the objects of consideration and reasoning patterns are important keys for semantic search.

Our mathematical ontology is a partial solution to this problem. The mathematical ontology has double layered structures in which both the classification by the study items of school textbooks and the contents of the learning items at the corresponding study items are separated. The ontology at study item structural level relates to school textbook’s classification, which is conventional, and the ontology at the
study content level relates to the contents of questions and related mathematical ideas, which is a solution to the contemporary demand.

The mathematical ontology is useful not only for teachers but also for learners. Despite many similar questions in essentials, average learners do not see their similarity just because the way of writing and the subject matters are different. This tool provides learners with an opportunity to think about the nature of questions and change their viewpoints of mathematical ideas.

In view of these, our approach flexibly responds to the education environment change.

CONCLUSION

We constructed mathematical ontology and developed the searching tool based on the ontology in the previous work. In this paper, we reported about a newly added function that allows the retrieval by the difficulty level of questions, mainly determined by the type and degree of a mathematical expression. We also developed a supporting tool for inputting questions with MathML.

Newly introduced retrieval patterns help teachers prepare for teaching materials in class. Our searching tool makes it possible to compile quality teaching materials not only because the tool reduces the burden of teachers but also the tool provides new dimensions in semantic searching patterns.

The difficulty levels of questions treated in this study are based on the criteria given by the experts of mathematical area or teachers’ empirical measure. However, we showed that, by analyzing the contents of mathematical expression denoted by MathML, we can add a new viewpoint for determining the difficulty level.

There still remain questions that are not dealt with by the tool. For example, questions about graphical or complex word problems. And there are other retrieval patterns desired in educational area. For this tool to be popular, we need to deal with more and more questions in both quantitative and qualitative terms.

In another respect, many teachers are very busy with preparations for class and various school events. So, for effective utilization of this kind of tool, the important point is as follows:

1. No technical knowledge other than mathematics is required.
2. No need to read instruction manuals.

We improve OMQS with particular emphasis on the above mentioned points.

We also plan to apply the result of our study to other fields where the mathematical expression is used after verifying the feasibility of the mathematical ontology and the searching tool.

REFERENCES