Bulletin of TICMI
Vol. 27, No. 1, 2023, 29–37

About Knowledge Delivery Strategies for Intelligent Tutoring Systems in Mathematics and Computer Science

Nikoloz Abzianidze[b*], Natela Dogonadze[b], Giorgi Ghlonti[a,b], Zurab Kipshidze[a]

[a]*Muskhelishvili Institute of Computational Mathematics*
Georgian Technical University
4 Grigol Peradze St., 0159 Tbilisi, Georgia;
[b]*International Black Sea University*
2, David Agmashenebeli Alley 13, 0159 Tbilisi, Georgia
(*Received February 2, 2023; Revised April 24, 2023; Accepted May 3, 2023*)

The paper aims to develop methods for increasing knowledge and model tracing capabilities of intelligent tutoring systems designed for teaching mathematics and informatics. From different areas of mathematics and computer science, based on the strategy of self-explanation, the authors consider the cases where the solution of a problem can be achieved as a result of generalization of results found at previous stages. In case of application of holistic approach and appropriate methods of ontology engineering this can become the basis for building an e-learning environment, when the student naturally moves through the subject based on effective feedback from the ITS.

Keywords: Cognitive tutors, model tracing; knowledge tracing, self-explanation strategy, holistic approach; knowledge engineering.

AMS Subject Classification: 97U50.

1. Introduction

Intelligent tutoring systems (ITS) are complex, coordinated computer program frameworks that apply the standards and strategies of artificial intelligence (AI) to the issues and needs of educating and learning. They are created on the premise of four interlinked components [1]:

- Domain model
- Learning model
- Students model
- User interface

The domain model represents a subject area of the study in some formalized way. The learning model is responsible for tutoring strategies and actions. The student model represents the learner in terms of his potential and realized opportunities, achievements and problems that arise in the process of mastering the subject. The user interface is responsible for communication. The knowledge accumulated and provided by the domain model may be regarded from two different perspectives - epistemological point of view and computational point of view.

* Corresponding author. Email: nabzianidze@ibsu.edu.ge

ISSN: 1512-0082 print

Epistemology refers to the nature of knowledge, value and validity of it, methodological considerations about construction of knowledge. Epistemology makes the foundation for the methods of formalization and inference, further realized in computational component of tutoring systems [2].
From computational perspective, intelligent tutoring systems have gone through the following stages of development:

- black box models;
- glass box models;
- cognitive models [3].

The ITS constructed according to the black box model is responsible only for providing the final result for the given input, not taking into account the mental activity of a student, simply checking whether the answer is true or false. If the ITS is constructed in accordance with the glass box model, it manipulates the same domain constructs as a human expert but does not simulate his way of reasoning. When the ITS is constructed according to the cognitive model, it seeks to match knowledge representation formalism and inference mechanisms with human cognition, trying to recreate the way knowledge is accumulated and manipulated by human thought, for the ability to assist the learner during the complete path of his/her problem-solving activities.
Based on these considerations, cognitive tutors are built on the basis of two complementary techniques - model tracing technique and knowledge tracing technique. Model tracing means ability of a tutoring system to have control over students progress through a problem solution. Knowledge tracing means ability of a system to have control over students learning from problem to problem. From the point of view of architecture, two loops can be distinguished in the structure of cognitive tutors - an outer loop, responsible for knowledge tracing and an inner loop, responsible for model tracing [4].
Whatever computational model would not be selected, there is a necessity in the formalism to describe the knowledge accumulated in a subject area. There are several ways of representing these models. Roger Nkambou in [2] mentions the following:

- production rules;
- semantic networks;
- conceptual graphs;
- frame-based models;
- ontologies and description logic.

A production rule system includes a rule base that contains a description of algorithmic steps applicable to all the problems that the system may be asked to solve. The interpreter responsible for solution determines which rule to execute on each step of the activity.
Semantic networks describe a subject area in the form of objects (or classes of objects) and relations among them. This form of knowledge representation is especially suitable for describing the taxonomic structure of categories for domain objects and for expressing general statements about the domain of interest.
The formalism of conceptual graphs is based on semantic network, but is more rigorous with the ability to represent semantics of a domain, and is directly linked to the language of the first-order predicate logic.

Frame-based systems are based on the notion of frames or classes, representing the collections of instances. Each frame has an associate collection of attributes which can be charged by values or some other frame (thus expressing the relations of inclusion and inheritance among classes).
Ontology is a formal specification of a subject area. It includes a definition of concepts and the relationships among them. Description Logics are able to provide ontologies with a solid basis for inference and reasoning. Standard formalism developed for describing ontologies makes it possible to share and reuse them between different environments.

2. Knowledge acquisition mechanisms

According to modern pedagogy, knowledge in any subject area can be divided into at least three categories [5]:

- factual knowledge;
- conceptual knowledge;
- procedural knowledge.

Factual knowledge is represented by information that must be learned through repetition and commitment to memory. Conceptual knowledge means understanding the concepts accumulated in the subject area and the relations between them [6]. Conceptual knowledge forms the basis for building categories and serves to organize knowledge in the subject area. Procedural knowledge refers to the knowledge of procedures and algorithms for certain actions. This also means the ability to justify and check the correctness of the procedure based on specific models and methods, along with the possibility of modifying and expanding them.
The relations between conceptual and procedural knowledge can be described in different theoretical schemes [7]. However, currently the most common is the so-called iterative concept, which implies the gradual deepening of each component over time, as it became clear that one type of knowledge cannot be formed completely without the formation of another, and that the acquisition of new knowledge requires the students active involvement in the process of constructing his own knowledge [8].
One additional component here might be added in the form of problem-solving skills as the ability of identification and formulation of problems, possibility of application of different models and strategies, and the ability to analyzing solutions. Another dimension to this categorization schema - the so-called cognitive dimension is added by Blooms taxonomy. According to the revised version of the taxonomy [9], cognitive processes by which thinkers encounter and work with knowledge may be represented as remembering, understanding, applying, analyzing, evaluating and creating.
When abbreviated multiplication formulas are taught in introductory algebra courses, they are usually presented in the form of the following identity transformation:

$$(a-b)(a+b)=a^2-b^2$$
$$(a+b)(a^2-ab+b^2)=a^3+b^3$$
$$(a-b)(a^2+ab+b^2)=a^3-b^3$$

This is understandable, since at the beginning, it is important to develop the skills that allow to apply these rules to simplify more complex expressions (on the basis of these formulas).
But further it is quite convenient to offer students in-depth analysis of these transformations. In particular, in case of a binomial $a^2 - b^2$, we may offer students to fill it with monomials of the degree 2 consisting from the variables a and b. There is only one such - ab.
Therefore, we get:

$$a^2 - b^2 = a^2 - ab + ab - b^2 = a(a - b) + b(a - b) = (a - b)(a + b)$$

This method is productive, as it can be easily generalized to third degree binomials. In this we'll have two monomials - a^2b and ab^2. Accordingly, we get:

$$\begin{gathered} a^3 - b^3 = a^3 - a^2b + a^2b - ab^2 + ab^2 - b^3 \\ = a^2(a - b) + ab(a - b) + b^2(a - b) = (a - b)(a^2 + ab + b^2) \end{gathered}$$

Eventually this method naturally leads to the formula for the decomposition of binomials of degree n:

$$\begin{gathered} a^n - b^n = a^n - a^{(n-1)}b + a^{(n-1)}b - a^{(n-2)}b^2 \\ +a^{(n-2)}b^2 - \cdots + a^2b^{(n-2)} - ab^{(n-1)} + ab^{(n-1)} - b^n \\ = a^{(n-1)}(a - b) + a^{(n-2)}b(a - b) + \cdots + ab^{(n-2)}(a - b) + b^{(n-1)}(a - b) \\ = (a - b)(a^{(n-1)} + a^{(n-2)}b + \cdots + ab^{(n-2)} + b^{(n-1)}) \end{gathered}$$

It is known that when it comes to mathematical concepts, at the initial stages, students traditionally perceive them as processes. For example, when moving from an arithmetic course to an algebra course the expression

$$y = 3x + 1$$

is usually considered as a procedural mechanism that allows the calculation of the value of the variable y based on the value of the variable x.
But it can be considered as well as a function that maps the set of values of the variable x to the set of values of the variable y. This is already a pattern of vision formed as a result of abstraction.
The same way, in the Theory of Relativity, the Lorentz transformation can be considered as a mechanism that allows to determine the value of coordinates and time point of an event in one reference system on the basis of those values in another reference system. But on the other hand, these transformations can be considered as mappings corresponding to rotations in Minkowski four-dimensional pseudo-Euclidean space.
One of the theories that helps us understand how people, not only in science, but also in real life, build abstract concepts based on concrete actions, is the so-called APOS (actions, processes, objects, schemas) theory [10]. According to the APOS theory, the first step from the concrete to the abstract lies in action - to understand mathematical concepts it is necessary first of all to perform appropriate actions and to apply certain rules.

After that, we will consider the process itself as an object on which certain manipulations can be performed. For example, $y = x^2$ can be perceived as a parabola (object) on which various processes, such as shift, stretch, compression, inverse of a function, and its composition with other functions are done.

The final step of abstraction is the construction of a schema. A schema is a set of diverse interpretations of concepts and/or constructs responsible for synthesis, generalization and retrieval of similar experiences [11]. In this regard, the history of the formation of the Theory of Groups by Evarist Galois gives us an interesting example.

3. Self-explanation as learning strategy

One of the mechanisms of active learning is the so-called self-explanation. This is a constructive activity that allows the student to be deeply involved in the teaching process and to monitor it effectively [12].
A number of cognitive mechanisms are involved in this process. Among them reasoning to fill the lack of information, integration of acquired information with learning material and existing knowledge, monitoring and correction of misconceptions, etc. The following programming task can be cited as an example of application of self-explanation as a learning strategy:

We are given a 3×3 matrix

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Problem is to write a program that calculates its determinant.

Solution: Let us calculate the determinant according to the following formula:

$$\begin{aligned} det(A) = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{21} \cdot \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \cdot \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ = a_{11} \cdot (a_{22}a_{33} - a_{23}a_{32}) \\ +a_{21} \cdot (a_{12}a_{33} - a_{13}a_{32}) \\ +a_{31} \cdot (a_{12}a_{23} - a_{13}a_{22}). \end{aligned}$$

If we look at resulting terms, we can see that they have the following structure:

$$a_{j1} \cdot (a_{i2}a_{k3} - a_{i3}a_{k2})$$

satisfying the following constraints:

$$j \in [1, 2, 3], i \in [1, 2, 3], k \in [1, 2, 3], j \neq i, j \neq k, i \neq k. \tag{1}$$

Therefore, the first task that can be given to students should be in formation of the structures

$$\{(j, [(i, k)])\}$$

in such a way that conditions (1) are fulfilled. This finally gives us the following output:

$$\{(1,[(2,3),(3,2)]);(2,[(1,3),(3,1)];(3,[(2,3),(3,2)])\}$$

Once this task is fulfilled, we can ask students to perform the following transformation:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \rightarrow$$
$$\rightarrow \{(a_{11},[a_{22}a_{33}-a_{23}a_{32}]),(a_{21},[a_{12}a_{33}-a_{13}a_{32}]),(a_{31},[a_{12}a_{23}-a_{13}a_{22}])\}.$$

Only then we ask them to perform the arithmetic operations and get the final result.

The use of self-explanation learning strategies can give interesting results in more complex cases as well. For example, consider the tensor transformations in the General Theory of Relativity.

As it is known, the starting point of Einsteins Theory of Gravitation is the equation of motion for a material point, recorded in free-falling, locally inertial reference frame:

$$\frac{d^2\xi^\alpha}{d\tau^2}=0$$

where $\alpha \in [0,1,2,3]$, and

$$d\tau^2=-\eta_{\alpha\beta}d\xi^\alpha d\xi^\beta$$

is a scalar remaining invariant to Lorentz transformation.

If we rewrite this equation for any other freely falling reference frame system, in the general case, it will take the following form:

$$\frac{d^2x^\lambda}{d\tau^2}=\Gamma_{\mu\nu^\lambda}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau},$$

where

$$\Gamma^\lambda_{\mu\nu}=\frac{\partial x^\lambda}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}$$

and $d\tau^2$ can be expressed by the following formula

$$d\tau^2=-g_{\alpha\beta}dx^\alpha dx^\beta,$$

where $g_{\alpha\beta}$ is the so-called metric tensor and

$$g_{\alpha\beta}=\frac{\partial x^\mu}{\partial\xi^\alpha}\frac{\partial x^\nu}{\partial\xi^\beta}\eta_{\mu\nu}$$

As for the structure $\Gamma^{\lambda}_{\mu\nu}$ it is called the affine connection.
The study of mathematical properties of the metric tensor and the affine connection is important for the analysis of gravity as a phenomenon.
First of all, let us note that

$$\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} = \Gamma^{\rho}_{\lambda\mu} g_{\rho\nu} + \Gamma^{\rho}_{\lambda\nu} g_{\rho\mu}. \tag{2}$$

It is more difficult to get the following result:

$$\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} + \frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\lambda}}{\partial x^{\nu}} = 2\Gamma^{\rho}_{\lambda\mu} g_{\rho\nu}. \tag{3}$$

Let us try to build a scheme that, based on transformation 2, will allow us to get this result, using the self-explanation strategy:
First of all, we note that the right-hand side of the transformation (2) includes two terms $\Gamma^{\rho}_{\lambda\mu} g_{\rho\nu}$ and $\Gamma^{\rho}_{\lambda\nu} g_{\rho\mu}$. Therefore, let us present it in the following form:

$$\Gamma^{\rho}_{\lambda\mu} g_{\rho\nu} + \Gamma^{\rho}_{\lambda\nu} g_{\rho\mu} = a + b. \tag{4}$$

After that we offer the following tasks to students:

(1) Since according to equation (1), $\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}}$ decomposes into the sum of two terms, let us change the μ, ν, λ indices to get an expression that can be interpreted as

$$a + c,$$

so that a coincides with the first term from equality 4, but $c \neq b$;
Such a result will be given by $\frac{\partial g_{\lambda\nu}}{\partial x^{\mu}}$.

(2) Let us use the equality

$$a + b + a + c - b - c = 2a$$

to get finally the following result

$$\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} + \frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\lambda}}{\partial x^{\nu}} = 2\Gamma^{\rho}_{\lambda\mu} g_{\rho\nu}.$$

Within the framework of the self-explanation strategy, another approach may be proposed to solve the above-mentioned problem:
Let us pay attention to the fact that both terms from the right side of the transformation (2) ($\Gamma^{\rho}_{\lambda\mu} g_{\rho\nu}$ and $\Gamma^{\rho}_{\lambda\nu} g_{\rho\mu}$), represent a product where the first multiplier is symmetric with respect to lower indices. The upper index is muted by one of the indices of the second multiplier, and in turn, this multiplier is also symmetric with respect to its indices.

To use a self-explanatory strategy, let us replace $\Gamma^{\rho}_{\lambda\mu}$ with a simpler, but somewhat analogous structure. The mentioned procedure of substitution does not imply, in any case, an identical transformation, but rather modeling. Instead of performing transformations on quite a complex structure, we perform the analysis of a simpler

one, in which the features important for the performance of the given transformation are outlined.
In particular, let us make the following substitution:

$$\Gamma^{\rho}_{\lambda\mu} g_{\rho\nu} \rightarrow T^{\rho} a_{\lambda} a_{\mu} b_{\rho} b_{\nu}$$

Here T^{ρ} represents the contravariance of the ρ index, and the other members represent (express) the covariance of the corresponding indices. In addition, the product $a_{\lambda} a_{\mu}$ ensures the symmetry of the initial expression with respect to the indices λ and μ, and the product $b_{\rho} b_{\nu}$ ensures its symmetry with respect to the indices ρ and ν. The only additional requirement is that the multipliers a and b do not commute with each other, but as we will see, this condition is not decisive in the case of our transformation. To simplify the situation even more, recall that ρ is a dumb index and make another substitution:

$$T^{\rho} a_{\lambda} a_{\mu} b_{\rho} b_{\nu} \rightarrow a_{\lambda} a_{\mu} b_{\nu}$$

Thus we come to the following chain of substitutions:

$$\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} + \frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\lambda}}{\partial x^{\nu}} \rightarrow$$
$$T^{\rho} a_{\lambda} a_{\mu} b_{\rho} b_{\nu} + T^{\rho} a_{\lambda} a_{\nu} b_{\rho} b_{\mu} \ + T^{\rho} a_{\mu} a_{\lambda} b_{\rho} b_{\nu} + T^{\rho} a_{\mu} a_{\nu} b_{\rho} b_{\lambda} T^{\rho} a_{\nu} a_{\mu} b_{\rho} b_{\lambda} \ - T^{\rho} a_{\nu} a_{\lambda} b_{\rho} b_{\mu}$$
$$\rightarrow a_{\lambda} a_{\mu} b_{\nu} + a_{\lambda} a_{\nu} b_{\mu} + a_{\mu} a_{\lambda} b_{\nu} + a_{\mu} a_{\nu} b_{\lambda} - a_{\nu} a_{\mu} b_{\lambda} - a_{\nu} a_{\lambda} b_{\mu}.$$

Next, we group similar members and perform the identical conversion in the last expression of the chain

$$a_{\lambda} a_{\mu} b_{\nu} + a_{\lambda} a_{\nu} b_{\mu} + a_{\mu} a_{\lambda} b_{\nu} + a_{\mu} a_{\nu} b_{\lambda} - a_{\nu} a_{\mu} b_{\lambda} - a_{\nu} a_{\lambda} b_{\mu},$$

which is practically a trivial result (and what is very important, nowhere has there been an attempt to interchange the coefficients a_i and b_j, which would be inadmissible due to their non-commutability). After developing these skills, performing conversion (3) will be much easier for students.

Conclusion

Above mentioned schemas for self-explanation strategies allow to strengthen model and knowledge tracing capabilities of cognitive ITSs for teaching mathematics and informatics. These techniques might be applied to a diverse set of courses in these subject areas. Naturally, this can be achieved under the condition of proper planning of knowledge delivery process, with application of holistic approach to knowledge engineering activities both at epistemological and computational level. The application of ontology engineering methods seems to be most promising.

Acknowledgements.

G. Ghlonti was partially supported by the European Commission HORIZON EUROPE WIDERA-2021-ACCESS-03 Grant Project GAIN (grant agreement no.101078950).

References

[1] P. Phobun. *Adaptive intelligent tutoring systems for e-learning systems*, Social and Behavioral Sciences, **38**, (2010), 4064-4069. doi: 10.1016/j.sbspro.2010.03.641

[2] R. Nkambou, J. Bourdeau, R. Mizoguchi (Eds). *Advances in Intelligent Tutoring Systems*, Studies in Computational Intelligence, Volume 308. Systems Research Institute. Polish Academy of Sciences. ISBN 978-3-642-14362-5. Springer-Verlag Berlin Heidelberg. 2010

[3] B. P. Woolf. *Building Intelligent Interactive Tutors*, Student-centered strategies for revolutionizing e-learning. ISBN: 978-0-12-373594-2. Elsevier Inc. 2009

[4] K. R. Koedinger, J. R. Anderson, W. H. Hadley, M. A. Mark. *Intelligent Tutoring Goes to School in Big City*, International Journal of Artificial Intelligence in Education. Issue 8. 1997.

[5] M.A.Al-Mutawah, T. Ruby, A. Eid, E.Y. Mahmoud, M.J. Fateel. *Conceptual understanding, procedural knowledge and problem-solving skills in mathematics: high school graduates work analysis and standpoints*. International Journal of Education in Practice, **8**, 3 (2019), 258-273, doi:10.1007/978-3-642-14363-2

[6] B. Rittle-Johnson, M. Schneider. *Developing conceptual and procedural knowledge in mathematics*, In R. Cohen Kadosh, A. Dowker (Eds.). Oxford Handbook of Numerical Cognition (pp. 1102-1118). Oxford UK: Oxford University Press. 2015. doi: 10.1093/oxfordhb/9780199642342.013.014

[7] N.M. Crooks, M.W. Alibali. *Defining and measuring conceptual knowledge in mathematics*. Development Review., **34** (2014), 344-377

[8] M.T.H. Chi. *Self-explaining expository texts: The dual processes of generating inferences and repairing mental models*. In R. Glasier (Ed.). Advances in Instructional Psychology, Mahwah, NJ: Lawrence Erlbaum Associates. 2000

[9] R. Clark, L.Chopeta. *Graphics for Learning: Proven Guidelines for Planning, Designing, and Evaluating Visuals in Training Materials*, San Francisco: Jossey-Bass/Pfeiffer. 2004

[10] E. Dubinsky. *Actions, Processes, Objects, Schemas (APOS) in Mathematics Education*, In Lerman, S. (eds) Encyclopedia of Mathematics Education. Springer, Dordrecht, https://doi.org/10.1007/978-94-007-4978-8-3. 2014

[11] Morehouse Kathryne Elisabeth. *Building conceptual understanding and algebraic reasoning in Algebra*, Education and Human Development Masters Thesis. The College at Brockport: State University of New York. 8.1.2007. $digitalcommons.brockport.edu/ehd-thesis/433$

[12] M. Roy, M.T.H. *The self-explanation principle in multimedia learning*, In R.E. Mayer (Ed.). Cambridge Handbook of Multimedia Learning, pp. 271-286. New York: Cambridge University Press. 2005. doi:10.1017/CBO9780511816819.018