Adventures in the Boundary between Domain-Independent Ontologies and Domain Content for CSCL

Seiji Isotani and Riichiro Mizoguchi

The Institute of Scientific and Industrial Research, Osaka University,
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan
isotani@acm.org, miz@ei.sanken.osaka-u.ac.jp

Abstract. One of the main problems facing the development of ontology-aware authoring systems (OAS) is to link well-designed domain-independent knowledge (ontologies) with domain content. Such a problem comes from the fact that all OAS developed to date require end-users (non-experts) to create their own domain ontologies to run the system in real scenarios. In collaborative learning (CL), this problem hinders the development of OAS that aid the design of pedagogically sound CL sessions with strong technological support. In this paper, we propose a framework that connects an ontology for CL (CL ontology) with domain content without the use of domain ontologies. To check its usability, we present an example to model a geometry drawing course demonstrating that it is feasible to instantiate the CL ontology to represent a specific domain and connect it with adequate learning objects.

Keywords: ontological engineering, CSCL, ontology-aware authoring system.

1 Introduction

Ontologies have been employed in many intelligent educational systems with some degree of success [3]. Although their use covers a vast field, ontologies are especially important to allow for a more explicit representation of knowledge in intelligent authoring systems (IAS) for education where the knowledge is based on various instructional/learning theories. One of the main problems facing the development of IAS for collaborative learning (CL) is to link well-designed domain-independent knowledge (ontologies) with the contents of a specific domain. On one hand, we have a very powerful and sharable knowledge that can be used in many different situations to support the authoring of CL sessions with theoretical justifications. On the other hand, we have domain content that needs to be adequately connected with theoretical foundations to support a well-designed CL session through the use of technology.

To solve the problem of connecting ontologies with specific domains and learning objects (LO), related researches [6;7] ask end-users to create their own ontologies for the specific domain. This approach show many benefits and good results for semantic annotation; however, both are excessively time-consuming due to the fact that they require end-users or non-experts to create their own ontologies from scratch.

This research proposes a framework to deal with this problem that connects domain-independent ontologies, specifically the CL ontology [9], with domain-specific content and LOs without the necessity of asking end-users to create new ontologies. This approach promotes a user-friendly way to implement the CL ontology by offering a framework along with templates that help users to link adequate LOs with the instantiated concepts in the ontology. In this paper we first show an overview of the prototype of an ontology-aware system that uses the CL ontology to support the design of CL sessions with theoretical justifications. Then we propose a framework for linking the CL ontology with domain-specific LOs. This framework is strongly based on our model GMIP that is presented in previous works. Finally, we demonstrate how to use this framework in a domain-dependent context.

2 Overview of CHOCOLATO

There are many learning theories that support group activities (*e.g.* [2;10;13]). Thus, to create group activities we can select an appropriate set of theories, considering necessary pre-conditions and desired benefits for learners. This flexibility of choosing suggests the difficulty of users (*e.g.* teachers) in selecting an appropriate set of learning theories to ensure learners' benefits. Therefore, to help these users, we need an elaborate system that considers different learning theories to support the CL design.

In CSCL, ontologies have been successfully applied to solve the problem of representing learning theories to support CL [9;11]. Furthermore, our previous research analyzed seven different theories (e.g. [2;4;10]) used to support CL activities to propose the Growth Model Improved by Interaction Patterns (GMIP) [8]. The GMIP is a graph model based on an ontological structure that describes an excerpt of learning theory. It represents, in a simplified way, the learner's knowledge acquisition and skill development processes, and explains the relationships between learning strategies, educational benefits and the interactions used to achieve these benefits.

The GMIP has twenty nodes that represent the levels of the learner's development at certain stages of learning. Each node is composed of two triangles. The upper-right triangle represents the stage of knowledge acquisition, while the lower-left triangle represents the stage of skill development. The nodes are linked with edges that show possible transitions between nodes in compliance with [1] and [12]. Using the GMIP graph, we show the benefits of a learning strategy by highlighting its path on the graph and associating each edge with interactions activities (top-right of Figure 1).

In order to develop a system to support the design of CL activities based on ontologies and our model, we have been developing CHOCOLATO (Concrete and Helpful Ontology-aware Collaborative Learning Authoring Tool) an ontology-aware system that uses ontologies developed in the Hozo Editor (http://www.hozo.jp) to provide its theoretical knowledge. And the sub-system called MARI allows the visualization of learning theories on the screen using the GMIP (Figure 1).

Using ontologies and the GMIP, MARI can select appropriate learning theories and strategies and suggest a consistent sequence of activities for learners in a group. The suggestions given by our system are guidelines that can be used to propose CL activities based on theories which (a) preserve the consistency of the CL process and (b) guarantee a suitable path to achieve desired benefits.

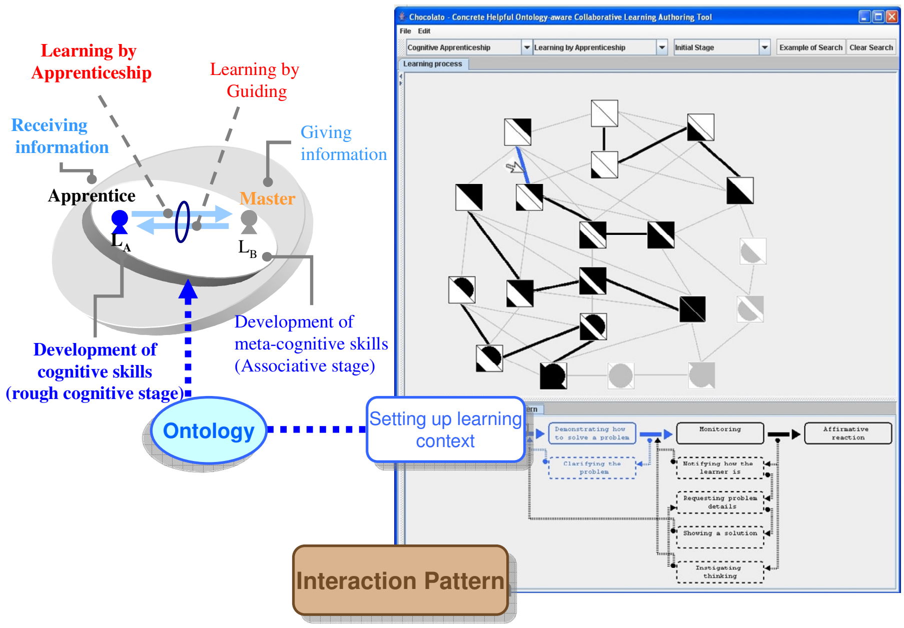


Figure 1: A screenshot of MARI showing the Cognitive Apprenticeship theory using the strategy *Learning by Apprenticeship*. The top shows the path on the GMIP. The bottom shows the sequence of activities in compliance with the theory. The left side shows other information that can be extracted from our ontology.

3 A Framework to Support Ontologies, Domain Content and LOs

MARI prototype is strongly based on domain-independent ontologies and our model GMIP. This means that MARI can provide domain-independent recommendations that can be used in different situations and are justified by theories, but it does not consider the actual domain (*e.g.* mathematics) in which the recommendations will be applied. Because of this, some colleagues/researchers have pointed out that although our approach is theoretically valid, it can hardly be applied in real environments. Thus, to augment our research and show that a theoretically valid approach can be applied in real environments, we propose a framework to link domain-specific content into our model GMIP and our ontologies. This framework is shown in Figure 2. The proposed framework has four linked layers. The top two layers are completely domain-independent, representing the knowledge about CL, learning theories and the learning state of a learner. The two bottom layers are related to domain-dependent content. One is related to the knowledge and skills of the domain-specific content and the other is related to the LOs connected with this content.

We define the learning state layer (top layer in Figure 2) as a set of nodes of different GMIPs. As summarized in section 2, each node in GMIP represents the stages of knowledge acquisition and skill development. Furthermore, each GMIP represents a different piece of knowledge and a different skill acquired by a learner. For example, to draw a geometric object a learner needs knowledge about the properties of the object and knowledge related to the manipulation of available drawing tools (*e.g.* square or compass). However, knowledge alone is not enough to draw a geometric object. A learner also needs the skill to use the properties of the geometric object correctly and the skill to adequately use the drawing tools. According to [1;9;12] we have four stages of knowledge acquisition: Nothing, Accretion, Tuning, Restructuring; and five stages of skills development: Nothing, Rough-cognitive, Explanatory-cognitive, Associative and Autonomous. In the learning state layer we represent each knowledge and skill being developed according to these stages.

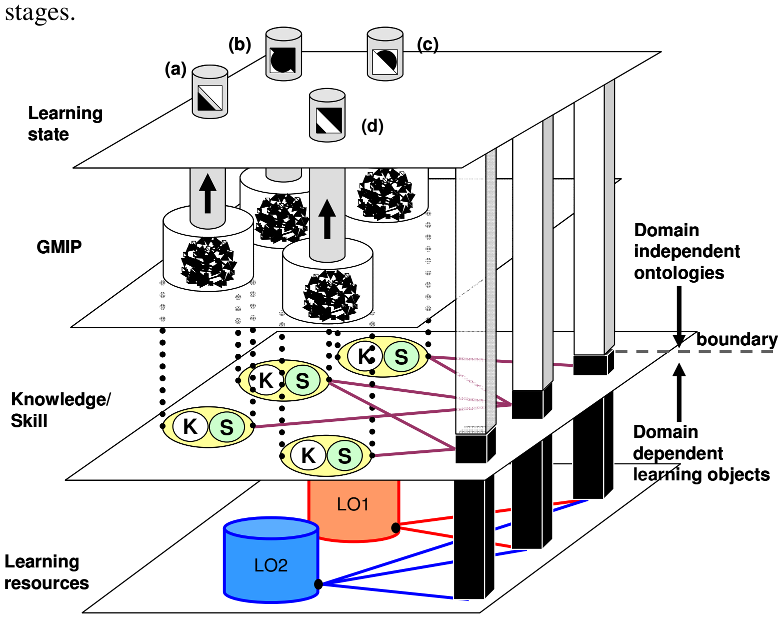


Figure 2. Framework to link domain independent ontologies, domain specific content and LOs.

The second layer is the GMIP. In this layer we show how learners can develop their knowledge/skills as transitions between nodes. In previous works we presented how this model was created and used to design CL activities. Now, the main difficulty is how to link the second and third layer in order to map the GMIPs with domain-specific knowledge and skills.

To define the third layer and link it with the second layer, first of all, given a domain-specific content and a learning goal, we must separate the knowledge from the skills necessary to achieve this goal in the specified domain. The knowledge to achieve the learning goal should be decomposed into different sub-knowledge pieces to be acquired. Similarly, the skills should be decomposed into sub-skills to be developed. The final structure will be a decomposition tree as shown in Figure 3 that identifies the minimum knowledge and skills necessary to construct a geometric object. The granularity of the decomposition tree depends on the learning goals and the expertise of the user who creates the tree. Observe that this tree is different from those proposed by [5, 11]. While those works provide decomposition trees that represent instructional design plans, our trees represent the knowledge and skills to be developed without any reference to how it will be developed. The design of CL activities can occur after linking this tree with the GMIP.

Using this approach, we can separate information about the content from information about how to learn the content. Such differentiation (boundary) is important when we think about learner-centered environments where the environments adapt the way to provide information or the way to teach the same content according to learning/teaching preferences.

To complete the mapping of knowledge and skills into our model GMIP it is necessary to explicitly identify the relationship of the knowledge and the skills in the tree as shown by the blue doted line in Figure 3. Each skill can be related to one or more pieces of knowledge and vice-versa. For each relationship between knowledge and skill we can create an instantiated GMIP, which will then be able to support the development of this knowledge and skill in the specific domain.

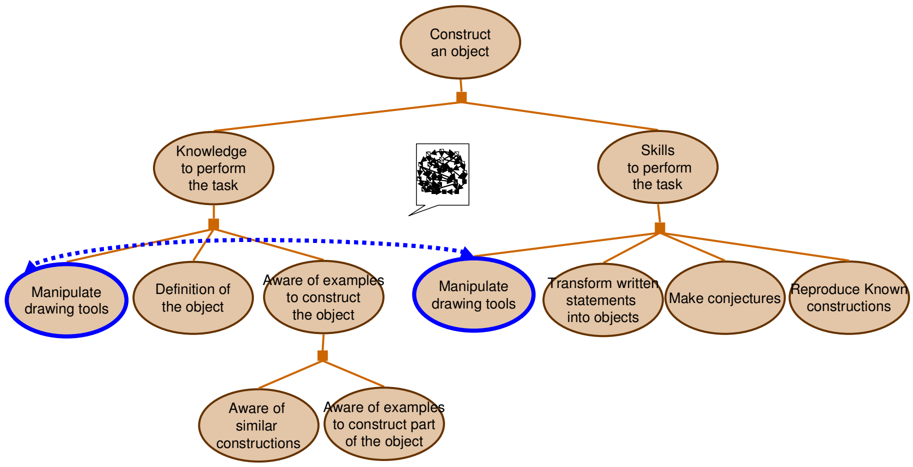
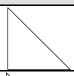
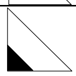
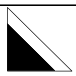
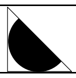


Figure 3. Example of a decomposition tree to separate knowledge and skills for a specific goal.

Finally, it is necessary to identify how each linked knowledge/skill fits into the stages proposed by [1; 9; 12]. To facilitate such a task, we provide templates that use the proposed definitions of stages of knowledge and skills. Such templates help users to adequately understand the knowledge and skill development process. Furthermore, it helps us to create a support system that semi-automatically maps specific skills/knowledge in our model GMIP. An example of the template for skill development instantiated for the example in Figure 3 (manipulate drawing tools) is shown in Table 1.

Table 1. Example of the template of skill development instantiated to represent the skill to manipulate a drawing tool. The phrases in bold are variables where specific skills can fit; all the others phrases come from the definitions provided by [1]. The GMIP graphical representation is shown in the middle column.

Stage Name	GMIP	Template definition
Nothing		Does not have the desired skills.
Rough-cognitive stage		Involves an initial encoding of **the skill to manipulate a drawing tool** into a form that is not sufficient for us to generate the desired behavior, usually by observing a process in which another person **manipulates a drawing tool**.
Explanatory-cognitive stage		Encodes **the skill to manipulate a drawing tool** into a form sufficient to permit a learner to **somewhat manipulate the drawing tool**.
Associative stage		Tunes the **skill to manipulate a drawing tool** through practice. Errors in the initial understanding of the skill are gradually detected and eliminated.
Autonomous stage		Demonstrates gradual continued improvements in the performance of the **skill to manipulate a drawing tool**.

The last layer in our framework is the learning resources layer. Each resource is a learning object (such as a tool, text, video, or activity) that can be used to improve a domain-specific knowledge or skill. Each learning object (LO) can be linked with different knowledge or skills. Thus, with the domain-specific goals provided in the third layer, an end user can add/remove LOs to satisfy specific conditions in the environment where the learning will occur.

This framework provides a high degree of flexibility to change the system. Thus, maintaining the same representation of knowledge and skills, changes in one layer do not affect the other layers. For example, a user can provide a new way to design a CL session, include this way as a path on GMIP, and the system can use it immediately without any change in the other layers. Similarly, a user can re-define the decomposition tree or add/remove LOs without any modification in the top layers. Because of this flexibility, it is possible to instantiate domain-independent CL ontologies to support the use of domain-specific content and LOs. Thus, we can offer much better support for users during the design of CL activities, taking into account the specific conditions of the domain and using the LOs available in the environment.

4 An Example of Application

To provide an example of application, we tested the usability of our framework to model a geometry drawing course. This course is one of the required courses for graduation in Mathematics at the University of Sao Paulo. It is comprised of two teachers, three teaching assistants and more than 150 students. In the geometry drawing course the flexibility of learning resources is very high. Thus, although the main learning goals remain the same, according to teachers' intentions, new exercises are included or removed from the curriculum while students are taking the course.

One of the main goals of this course is to provide learners with the knowledge and skills to (1) construct objects (such as an angle, a triangle, and a circle); (2) define objects (e.g., what is a point, an angle, a triangle); (3) classify objects according to some properties (e.g., triangle: equilateral or isosceles); (4) check properties (distances, angles, sides); (5) propose conjectures (e.g., this object is equal to another one because …); and (6) generalize and make inferences about properties (e.g., any square is a rectangle, but not vice-versa).

Using our framework, we created decomposition trees that explicitly identify the knowledge and skills that need to be acquired to achieve the objectives listed above, as shown in Figure 3 for objective (1). Then, using the templates we provided (Table 1), we mapped the knowledge and skills into our model GMIP. Finally, about 70 exercises were created and linked with the respective knowledge and skills to serve as learning objects for this trial.

To give an example of how our system for CL design will work based on this framework, let us propose a CL session that helps a learner to acquire the skills to construct a geometric object. To prepare such a CL session, using the MARI interface the user can select the initial stages (pre-conditions) of learners. This step corresponds to identifying the stages of the learners in the top layer. Because each stage is instantiated for a domain specific goal (Table 1), the system will be able to support the user by offering explanations of each stage of the learning development considering the specific domain. Thus, it can facilitate a more accurate selection of initial stages for learners. Furthermore, this process can be done semi-automatically.

Next, the user can select the desired final stage (desired goal) for learners. Through the formal structure based on ontologies which allows MARI to evaluate theories and their features, it will check the learners' conditions, learning goals and special requirements related to the domain to offer a sophisticated group formation and design CL activities with theoretical justifications. If more than one theory can help a group of learners, the system will present all of them and the user can select the preferable one. Our approach uses theory-driven group formation with suggestions of role assignment and sequence of interactions to offer fundamental settings for an effective CL session and essential conditions to predict the impact of interactions in the learning process. Furthermore, MARI can suggest different interaction patterns (sequences of activities) to achieve the desired goal. One example of the pattern for the Cognitive Apprenticeship Theory is shown in the bottom-right of Figure 1. This process is done on the second top layer of our framework. Due to the limitation of space we cannot explain how the group is formed or how the role assignment is done.

In the third layer, the system will take the recommendations of the second layer and use the decomposition tree with the information of the specific domain to support the user in implementing the CL session. For example, for the first three interactions of the interaction pattern in the bottom-right of Figure 1, we show on Table 2 the differences between domain-independent events and domain-specific events. *Master* and *Apprentice* are the roles that learners will play during the CL session. These interactions are proposed to help the apprentice to develop the skill to manipulate a drawing tool (from *Nothing* to the *Rough-cognitive stage*) and to help the master to develop his cognitive and meta-cognitive skills (from the *Associative stage* to the *Autonomous stage*).

Table 2. Three interactions from the Cognitive Apprenticeship interaction pattern showing the differences between domain-independent and domain-dependent activities to develop the skill to manipulate a drawing tool.

Interaction	Domain-independent activity		Domain-dependent activity	
	Master	Apprentice	Master	Apprentice
Set up learning context	Giving Information	Receiving Information	Showing the drawing tools that will be used and explaining their functionality	Becoming familiar with the drawing tools
Demonstrate how to solve a problem	Demonstration	Observing demonstration	Drawing one object using adequate tools and explaining the process	Observing the steps to draw the object using the tools
Clarify the problem	Identifying misconceptions	Externalization of misconceptions	Answering questions to identify weak points in the explanation and in the apprentice's skill in using the tools.	Asking questions about the correct way to use the tools.

Finally, each interaction of a pattern is linked with specific knowledge/skills in the decomposition tree. The knowledge/skills are then linked with specific learning objects that will support carrying out the CL process in the specific domain (fourth layer). In our example, in which the domain is geometric drawing, each interaction is associated with geometry exercises or texts explaining geometric concepts. Thus, the system can help the user to select adequate material to support the CL session and run group activities. For example, to support learners to manipulate drawing tools, learning objects related with the simple drawing of triangles, parallels, and circumferences are connected. In this case, drawing simple geometric objects is fundamental to familiarize learners with the given drawing tools.

This example demonstrates that it is possible to connect domain-independent ontologies with domain-dependent LOs in CSCL environments without asking end-users to create ontologies by themselves. Furthermore, because our ontologies are based on theories, our framework also gives some hints about how to use a theory to support real environments in a CL context. The next version of CHOCOLATO will have the functionalities presented in this section.

5 Conclusions

To create intelligent educational systems based on well-grounded theoretical knowledge and to apply them in real environments are two important challenges that research in the development of ontology-aware systems are facing nowadays. In order to solve these problems in the context of CL, we propose a framework that intends to connect CL ontology [9] with LOs intermediated by our model GMIP. By providing this connection, we can offer a more user-friendly way to design pedagogically sound CL sessions in a specific domain with strong technological support.

The proposed framework is divided into four layers interconnected by the concept of knowledge acquisition and skill development proposed by [1;9;12]. Thus, each layer has the flexibility to change any of its content or characteristics without affecting the other layers or system functionality. The top two layers (Figure 2) represent the domain-independent knowledge and the bottom two layers represent the domain-dependent knowledge that instantiates the top layers. To create the domain-dependent knowledge, a user needs to create a decomposition tree for the specific domain and map it into our model GMIP. This process is partially supported by our templates, which can be generated semi-automatically during the process of mapping. Such an approach seems to be more reliable than other approaches, especially because it removes the burden of creating domain ontologies for each domain of application.

To exemplify the use of our framework, we mapped a geometric drawing course (together with its LOs) into our model GMIP. This example demonstrated that it was feasible to instantiate the GMIP (and thus, the CL ontology) to represent the specific domain and connect it with domain-dependent LOs. Our future research intends to complete the implementation of CHOCOLATO using this framework in order to help users to (a) form groups and design CL activities in compliance with theories by suggesting adequate LOs during this process; and (b) analyze interaction among learners to identify the educational benefits acquired during the CL process.

References

1. Anderson, J. R. (1982) Acquisition of Cognitive Skill, Psychological Review, 89(4), 369-406.
2. Collins, A. (1991) "Cognitive apprenticeship and instructional technology". In: Idol, L., and Jones,B.F. (Eds.) Educational values and cognitive instruction., LEA..
3. Devedzic, V. (2006) Semantic Web and Education, Springer.
4. Endlsey, W. R. (1980) Peer tutorial instruction, Educational Technology.
5. Gagne, R., Briggs, L. & Wager, W. (1992). Principles of Instructional Design (4th Ed.). Fort Worth, TX: HBJ College Publishers.
6. Gasevic, D., Jovanovic, J. & Devedzic, V. (2005) Ontologies for Reusing Learning Object Content. In Proceedings of the Int. Workshop on Ontologies and Semantic Web for E-Learning. 650-659.
7. Knight, C., Gasevic, D., & Richards, G. (2005). Ontologies to integrate learning design and learning content. Journal of Interactive Media in Education, 2005(07), 1-24.
8. Isotani, S., Mizoguchi, R. (2007) Deployment of Ontologies for an Effective Design of Collaborative Learning Scenarios. In Proceedings of CRIWG, LNCS 4715, 223-238.
9. Inaba, A., Ikeda, M. & Mizoguchi, R. (2003) What Learning Patterns are Effective for a Learner's Growth?. In Proc. of International conference on Artificial Intelligence in Education, 219-226.
10. Lave, J., Wenger, E. (1991) Situated Learning: Legitimate peripheral participation, Cambridge University Press.
11. Mizoguchi, R. Hayashi, Y., & Bourdeau, J. (2007) Inside Theory-Aware and Standards-Compliant Authoring System. In Proceedings of the Int. Workshop on Ontologies and Semantic Web for E-Learning, 1-18.
12. Rumelhart, D.E., & Norman, D.A. (1978) Accretion, Tuning, and Restructuring: Modes of Learning. Semantic factors in cognition. LEA, pp. 37-53.
13. Salomon G. (1996) Distributed Cognitions: Psychological and Educational Considerations, Cambridge University Press.