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

Abstract. One of the main problems facing the development of ontology-aware authoring systems (OAS) is to link a well designed domain independent knowledge (ontologies) with content related to a specific domain. Such problem comes from the fact that all ontology-aware system developed to date requires end-users or non-expert users to create their own domain ontologies to run the system in real scenarios. In collaborative learning (CL), the problem of linking ontologies with domain-specific contents hinders the development of OAS that aids the design of pedagogically sound CL sessions with strong support of technologies. In this paper, we propose an alternative framework that intends to connect an ontology that represents CL explicitly, referred to as CL ontology, with domain-specific contents without the use of domain ontologies. To check its usability, we show an example of its application to model a geometry drawing course. In this example we demonstrated that it was feasible to instantiate the CL ontology to represent a specific domain and connect it with the adequate learning objects.

1 Introduction

Ontologies have been employed in many intelligent educational systems with some degree of success [4]. Although, its use covers a vast field, ontologies are especially important to allow for a more explicit representation of knowledge used by intelligent authoring systems (IAS). In IAS for Education, the representation of knowledge is quite challenging due to the fact that current knowledge concerning IAS can be based in various instructional/learning theories.

In collaborative learning (CL) these theories are particularly complex because of the context of group learning where the synergy among learner's interactions affects the learning processes and hence learning outcomes [16]. In this context, Inaba and colleagues conducted an ontology-based research on modeling CSCL for years to support the development of IAS for CL, strongly based on learning theories [9;10]. The efforts to enrich this ontology up to date have produced a large and heavy-weight ontology, referred to as collaborative learning ontology, which deals with many difficult problems for proposing a well thought out CL session.

To go a step forward in previous achievements on ontology research, one of the main problems facing the development of IAS for CL is to link a well designed domain independent knowledge (ontologies) with 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, there are domain-specific contents that need to be adequately connected with theoretical foundations to support a well designed CL session through the use of technology (including learning objects). In ontology-aware systems, both tools and knowledge are somewhat volatile. Thus, the systems do not depend on one set of tools or a static domain-specific knowledge structure.

To solve the problem of connecting ontologies with specific domains and learning Objects (LO), it was proposed in [7] the use of two kinds of ontologies: content structure ontologies and domain ontologies that, together with the knowledge of the system, allow the linking of domain dependent content and LOs. Another related work proposes the concept of learning object context (referred to as LOCO-Cite) to integrate learning designs and LOs [8]. These approaches show many benefits and good results for semantic annotation and reusability of LOs, however both are excessively time consuming due to the fact it requires end-users or non-experts to create their own ontologies from scratch.

In order to deal with this problem this research proposes a framework that connects domain independent ontologies, specifically the CL ontology, 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 use the CL ontology by offering a graphical visualization of information and templates that help users to link adequate LOs with the instantiated concepts in the ontology.

In this paper, initially, we show an overview of the prototype of an ontology-aware system that use the CL ontology and a model based on it (GMIP) to support the design of CL session with theoretical justifications. Then, we proposed the framework for linking the CL ontology with domain-specific LOs. This framework is strongly based in our model GMIP that were present in previous works. Finally, we demonstrate how to use this framework in a domain dependent context offering an example of linking the CL ontology with geometry exercises that have been used in an obligatory course for undergraduate students at the Institute of Mathematics and Statistics of University of Sao Paulo.

2 Overview of CHOCOLATO

To support group activities there are many learning theories (such as Anchored Instruction, Peer Tutoring, LPP, etc). Thus, to assign roles and strategies for learners in a group we can select appropriate set of learning theories considering necessary pre-conditions and desired educational benefits for learners. This flexibility of choosing different learning theories provides us with many ways to design and conduct learning processes. However, it also suggests the difficulty of selecting an appropriate set of learning theories to ensure learners' benefits and the consistency of the learning processes. Therefore, to help users (instructors, teachers, designers, etc) to design effective group activities we need an elaborated system that considers different learning theories to support the design in compliance with them.

The use of ontological engineering for knowledge systematization has shown significant results concerning how to represent the knowledge of educational environments considering theories [12]. In CSCL research, ontologies have been successfully applied to solve the problem of representing learning theories to support CL [9;10]. Nevertheless, there are some limitations: (a) it is not easy to determine which learning theory is appropriate to explain and support learner's development; and (b) it is difficult to propose activities in compliant with the theories to enhance interactions among learners and lead them to achieve desired goals.

To overcome these limitations, in previous work, we re-analyzed seven different learning theories frequently used to support CSCL activities, and then, we proposed the ***Growth Model Improved by Interaction Patterns*** **(GMIP)**. The GMIP is a graph model based on an ontological structure to describe an excerpt of learning theory. It represents, in a simplified way, the learner's knowledge acquisition process and skill development process, and explains the relationships between learning strategies, educational benefits and interactions used to achieve these benefits.

The GMIP graph has twenty nodes, which represent the levels of the learner's development at a certain stage 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 [13]. Using the GMIP graph, we show the benefits of learning strategies by highlighting its path on the graph and associating each edge with interactions activities (top-right of Figure 1).

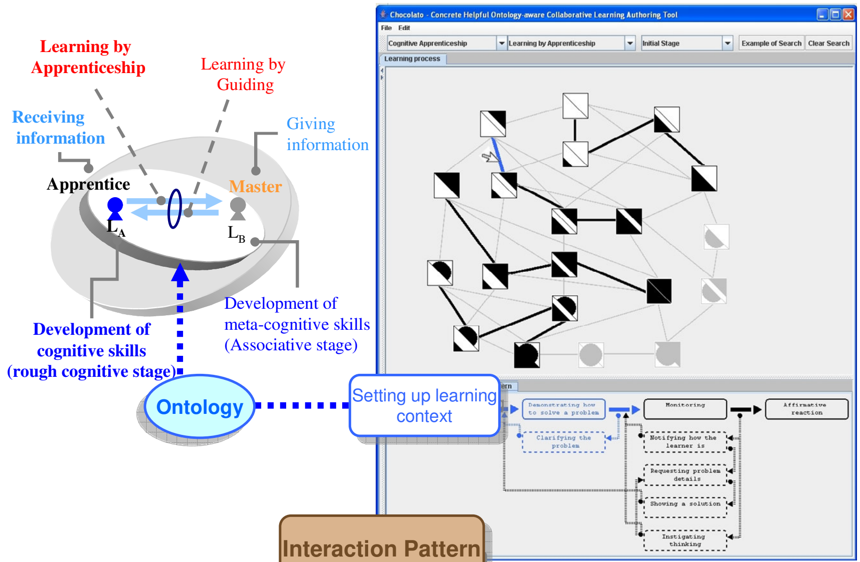


Figure 1. A screenshot of MARI showing the Cognitive Apprenticeship theory using the strategy *Learning by Apprenticeship*. On the top it shows the path on the GMIP and on the bottom the sequence of activities in compliance with the theory. On the left side, it shows some information that can be extracted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

In order to develop a system to support the design of CL activities, based on ontologies and our model, we have been developing **CHOCOLATO** – *a Concrete and Helpful Ontology-aware Collaborative Learning Authoring Tool*. It is an ontology-aware system that uses ontologies developed in Hozo ontology editor (http://www.hozo.jp) to provide its theoretical knowledge. One of its sub-systems called **MARI** – *Main Adaptive Representation Interface* allows to represent learning theories on the screen using the GMIP (Figure 1). Since 2007, MARI has been implemented in Java and uses Hozo API to interpret ontologies. Through the use of ontologies it allows high expressiveness and interoperability among theories and their features. Currently, it has in its database 6 theories: Cognitive Apprenticeship [2], Anchored Instruction [3], Peer Tutoring [5], Cognitive Flexibility [15], LPP [11] and Distributed Cognition [14].

Using ontologies and the GMIP, MARI can reason on the theories to select appropriate learning theories and strategies, and thus, suggest consistent sequence of activities for learners in a group. The **suggestions,** given by our system, are only guidelines for users 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. Using MARI, expert designers can propose their own sequence of activities. In such a case the system also can assist these users by providing other information that can be useful in various situations.

3 A Framework to Support Ontologies, Domain Content and LOs

Our prototype MARI is strongly based on domain-independent ontologies and our model GMIP. It means that MARI can provide domain independent recommendations, which can be used in many different environments and are justified by theories, but it does not consider the real domain (*e.g.* mathematics) in which the recommendations will be applied. Because of that, some colleagues/researchers have pointed out that although our approach is theoretically valid, it hardly can be applied in real environments. Thus, In order 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 last two bottom layers are related to domain-dependent content. One is related with the knowledge and skills about the domain-specific content and the other one is related to the LOs that can be connected with this content.

We define the learning state layer (top layer in Figure 2) as a set of nodes of different GMIP's. As summarized in section 2, each node in GMIP represents the stages of knowledge acquisition and those of skill development. Furthermore, each GMIP represents a different knowledge and a different skill acquired by a learner. For example, to draw a geometric object a learner needs the knowledge about the properties of the object and also needs the knowledge related with the manipulation of available drawing tools (pencil, square, compass, etc). However, only knowledge is not enough to draw a geometric object. Thus, a learner needs skills to use the properties of the geometric object correctly as same as skills to use adequately the drawing tools. According to [1;10;13] we have 4 stages of knowledge acquisition: Nothing, Accretion, Tuning, restructuring; and 5 stages of skills development: Nothing, rough-cognitive, explanatory-cognitive, associative and autonomous. In this layer we represent each knowledge and skill being developed according with these stages.

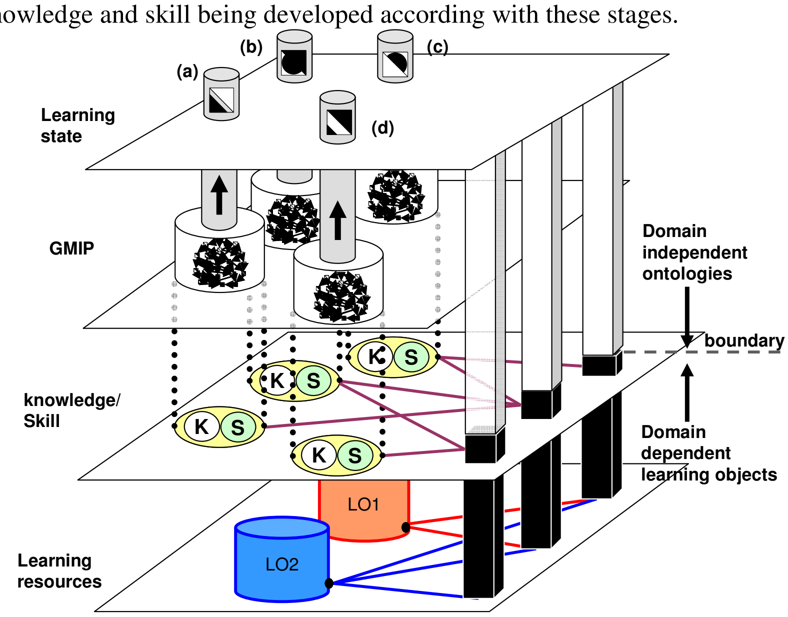


Figure 2: Framework to link domain independent ontologies, domain specific content and LOs.

The second layer is the GMIP. In this layer we show how a learner can develop his/her 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 linked with the second layer, first of all, given a domain-specific content and a learning goal, it is necessary to separate the necessary knowledge from the necessary skills to achieve this goal in the specified domain. The knowledge to achieve the learning goal should be decomposed into different sub-knowledge 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 in which it is shown a decomposition tree to identify 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 [6, 12]. While those works provide decomposition trees that represent the instructional design plans, our trees represent the knowledge and skills to be developed without any reference about how it will be developed. The design of CL activities can be acquired after linking this tree with the GMIP.

With this approach, we can separate the information about the content from the 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.

The next step 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 the blue doted line in Figure 3. Each skill can be related with one or more knowledge and vice-versa. For each relationship between knowledge and skill we can create an instantiated GMIP. Then, this GMIP will be able to support the development of this knowledge and skill in the specific domain.

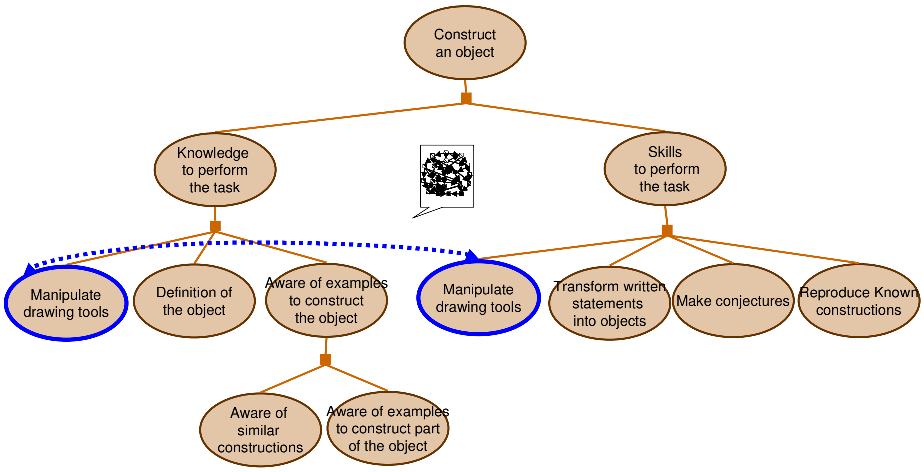
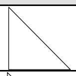
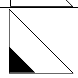
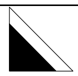
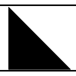
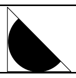


Figure 3: Example of decomposition tree to separate knowledge and skills for a specific goal.

Finally, for each linked knowledge/skill, it is necessary to identify how they fit into the stages proposed by [1;10;13]. To facilitate such 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 also 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 to fit in specific skills, all the others phrases come from the definition provided by [1].

Stage Name	GMIP representation	Template definition
Nothing		Does not have the desired skills.
rough-cognitive stage		It involves an initial encoding of **the skill to manipulate a drawing tool** into a form that it is not sufficient for us to generate the desired behavior, usually by observing a process in which other person **manipulate a drawing tool**.
explanatory-cognitive stage		It encodes **the skill to manipulate a drawing tool** into a form sufficient to permit a learner to **somewhat manipulate the drawing tool**.
associative stage		It is to tune the **skill to manipulate a drawing tool** through practice. Errors in the initial understanding of the skill are

		gradually detected and eliminated.
autonomous stage		It is one of the 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 (tool, text, video, activity, etc) 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 does 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. As same as before, a user can re-define the decomposition tree or add/remove LOs without any other modification in the top layers. Because of this flexibility, it is possible to instantiate domain independent CL ontology 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 of the framework presented in section 3, we tested its usability for modeling a geometry drawing course. The course of Geometry drawing is one of the obligatory courses to graduate in Mathematics at the University of Sao Paulo. This course comprised of two teachers, three teaching assistants and more than 150 students, divided into three groups. In this course the flexibility of learning resources (tools, exercises, etc) is very high. Thus, although the main learning goals remain the same, according to teachers' intention 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 (an angle, a triangle, a circle, …);
② Define objects (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 (this object is equal to another one because …);
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 for the topics listed above as shown in Figure 3 for the topic (1). Then, using the templates we provided (Table 1), we mapped these knowledge and skills into our model GMIP. Finally, about 70 exercises were created and linked with their 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 skills to construct a geometric object. To prepare such CL session, through the authoring interface of MARI using the GMIP, the user can select the initial stages (pre-conditions) of learners. This step corresponds to identify the stages of learner 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 explanation of each stage of the learning development considering the specific domain. Thus, it can facilitate a more accurate selection of initial stages for learners. Initially, the selection of stages will be done manually, but in the future 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 reasoning about the theories and its features (actions, roles, strategies, etc.), 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/strategy can help a group of learners, the system will present all of them and the user can select 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 (sequence of activities) to achieve the desired goal based on theories. One example of interaction pattern is shown in the bottom-right of Figure 1 for the Cognitive Apprenticeship Theory. All this process is done on the second top layer of our framework. Due to the limitation of space we cannot go deeper to explain how the group is formed and or how the role assignment is done.

In the third layer, the system will get the recommendations of the second layer and use the decomposition tree with the information of the specific domain to support the user to implement 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 rough-cognitive stage) and to help the Master to develop his cognitive and meta-cognitive skills (from associative stage to 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**
Setting up learning context	Giving Information	Receiving Information	Showing the drawing tools that will be used explaining their functionality	Familiarizing with the drawing tools
Demonstrate how to solve a	Demonstration	Observing demonstration	Drawing one object using the adequate tools	Observing the steps to draw the object

problem			explaining the process	using the tools
Clarify the problem	Identifying misconception	Externalization of a misconception	Answering questions to identify weak point in his own explanation and identify weak point 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. These knowledge/skills are linked with specific learning objects that will support to carry out the CL process in the specific domain (Fourth layer). In our example, in which the domain is geometry 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, for the interactions presented in Table 2, many exercises and texts related with simple drawings of triangles, parallels, and circumferences are connected. In this case, drawing simple geometric objects are fundamental to familiarize learners with the given drawing tools.

This example tries to demonstrate that it is possible to connect domain independent ontologies with domain dependent LOs in CSCL environments without ask an end-user (non-expert in building ontologies) to create ontologies by himself. 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 CL context. The next version of CHOCOLATO and MARI will have the functionalities presented in this section.

5 Conclusions

To create intelligent educational systems based on a well grounded theoretical knowledge and to apply it in real environments are two important challenges that research on the development of ontology-aware systems are facing nowadays. In order to solve this problem in the context of CSCL, we propose a framework that intends to connect the 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 session, in a specific domain, with strong support of technologies.

The proposed framework is divided in four layers interconnected by the concept of knowledge acquisition and skill development proposed by [1;10;13]. Thus, each layer has flexibility to change any of its content or characteristics without affect the other layers or the system functionality. The top two layers (Figure 2) represent the domain independent knowledge of the system and the bottom two layers represent the domain dependent knowledge that instantiate the top layers. To create the domain dependent knowledge a user need to create a decomposition tree for the specific domain and map it into our model GMIP. This process is partially supported by our templates that can be generated semi-automatically during the process of mapping. Such 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 tested its usability to map a geometry drawing course (together with its LOs) into our model GMIP. In this example we 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 researches 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, suggesting adequately LOs during this process; (b) analyze interaction among learners in order to identify the educational benefits acquired during the CL process; and (c) use the results of this analysis in order to improve the next CL sessions.

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. Cognition and Technology Group at Vanderbilt (1992) "Anchored instruction in science education" In: R. Duschl & R. Hamilton (Eds.), Philosophy of science, cognitive psychology, and educational theory and practice, Albany, NY: SUNY Press, pp. 244-273.
4. Devedzic, V. (2006) Semantic Web and Education, Springer.
5. 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.
8. 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.
9. Inaba, A., Supnithi, T., Ikeda, M., & Mizoguchi, R. (2000) How Can We Form Effective Collaborative Learning Groups. In Proc. of Intelligent Tutoring Systems, Springer, 282-291.
10. 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.
11. Lave, J., Wenger, E. (1991) Situated Learning: Legitimate peripheral participation, Cambridge University Press.
12. 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.
13. Rumelhart, D.E., & Norman, D.A. (1978) Accretion, Tuning, and Restructuring: Modes of Learning. Semantic factors in cognition. LEA, pp. 37-53.
14. Salomon G. (1996) Distributed Cognitions: Psychological and Educational Considerations, Cambridge University Press.
15. Spiro, R.J., Coulson, R.L., Feltovich, P.J., & Anderson, D.K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In The tenth annual conference of the cognitive science society. Hillsdale, NJ: Lawrence Erlbaum Associates, pp 375-383.
16. Stahl, G. (2005) Group Cognition: Computer Support for Collaborative Knowledge Building. Cambridge, MA: MIT Press.