Semantic Web-based Software Product Line for Building Intelligent Tutoring Systems

Alan Pedro da Silva¹, Evandro de Barros Costa², Ig Ibert Bittencourt³, Patrick H. S. Brito⁴, Olavo Holanda⁵, and Diego Demerval⁶

¹, ², ³, ⁴, ⁵, ⁶ GrOW - Group of Optimization of the Web
www.grow.ic.ufal.br

¹ Federal Institute of Education, Science and Technology
Palmeira dos Índios - Alagoas - Brazil
alan.silva@ifal.edu.br

², ³, ⁴, ⁵, ⁶ Federal University of Alagoas
{evandro, ig.ibert, patrick, olavoholanda, diego}@ic.ifal.br

Abstract. Intelligent Tutoring Systems (ITS) have been assumed as an important learning resource to be added as a module in e-learning systems. However, the construction of such systems is still a hard and complex task that involves, for instance, representation and manipulation of different knowledge source. To alleviate these issues, this paper proposes a new approach for building ITS by integrating Software Product Line and Semantic Web technologies focusing on two software engineering aspects: large-scale production and customization for different learners, and how to allow these knowledge to be automatically shared between software and authors in both reuse and knowledge evolution points of view. This paper shows a modeling for the proposed product line, as well as how the Semantic Web technologies was used to achieve the effective shared knowledge.

1 Introduction

Intelligent Tutoring Systems (ITS) are a kind of complex software system that can be used in an efficient way to improve the student learning process [1,11]. Such systems are provided with mechanisms that enable, in an automated fashion, a better understanding of the students’ needs, responding them individually. Moreover, ITS allow student centered learning despite the teacher centered learning, thus allowing the student to control his/her own learning process. In addition, ITS have been assumed as an important learning resource to be added as a module in e-learning systems.

However, the process of constructing Intelligent Tutoring Systems (ITS) is still a time-consuming, hard and complex task, and often requires a high cost of production. For instance, in order to obtain one hour of tutoring, about 200-300 hours of development must be spent [1]. This happens, especially, because an ITS has embedded Artificial Intelligence techniques. First, the complexity appears through the fact of many of these techniques can be used to represent the knowledge to be taught (Domain Model), to represent the knowledge of a student (Student Model), to reproduce the strategies of conventional teaching from a teacher, to implement new pedagogical strategies (Teaching Model), to
follow the learning and affective aspects (Diagnosis Process), and to motivate the student and enable an adequate interaction between the system and the student. Therefore, the project of an ITS demands a previous and very accurate study of how the techniques will be used, as the same Artificial Intelligence technique can be used for more than one purpose in an ITS as well as specific functionality can be implemented through integrating more than one AI technique. For instance, a mechanism for evaluating the knowledge of a student can be implemented through the integration of bayesian networks and rule-based reasoning, where each technique can offer different and complementary information about the student.

Actually, beyond the natural problematics of building a particular ITS, there are further limitations to increase even more the development process of such systems, especially if compared to traditional software [15,17,8,4]. The first problem is that every ITS is built independently, which makes any attempt of reusing any part of the system, especially the core components (Student Model, Domain Model, Diagnosis Process and Teaching Model) completely inviable. Another problem is the high maintenance cost due to the use of academics to develop such systems, by other means, using a development team which is not necessarily concerned about aspects of evolution and expansion.

To alleviate the mentioned issues, this paper proposes a new approach for building ITS by integrating Software Product Line and Semantic Web technologies focusing on two software engineering aspects: large-scale production and customization for different learners, and how to allow these knowledge to be automatically shared between software and authors in both reuse and knowledge evolution points of view.

According to [5], a Software Product Line (SPL) is a set of software-intensive systems sharing a common, managed set of features that satisfy the specific needs of a particular market segment or mission and that are developed from a common set of core assets in a prescribed way. Based on this definition, SPL offers good conditions for its applicability on the particular context of Intelligent Tutoring Systems because various tutors, in different domains, have common and variable features (functionalities) [6], which allow the creation of an unified platform for tutoring systems that can be adapted according to the individual needs of each author. Moreover, when it comes to software reuse, Software Product Line is one of the most sophisticated concepts in software engineering, which is focused on small-grained, opportunistic and technology-driven software reuse [13]. The strategic reuse provided by SPL combines the business strategies of the market segment (ITS in this work) with the technical strategies of reuse.

Therefore, SPL is an adequate approach for providing software reuse. However, traditional product line approaches do not address questions regarding knowledge reuse and knowledge sharing or evolution. It means an important lack in terms of ITS development and lifecycle. In this sense, the role of the knowledge engineer becomes very important because the correct implementation of an ITS is linked to how well the knowledge is specified. In addition, an ITS deal with knowledge of various types, making necessary the existence of a broad knowledge about the relationship between them, so they can be joint with the purpose of providing learning. Thus, in the process of building an ITS, it is necessary to first express the knowledge, then perform the implementation. This approach becomes inefficient in systems such as Intelligent Tutoring Systems,
since the number of knowledge domains an ITS can work is great.

This paper proposes a Product Line approach based on Semantic Web technologies. This is important for the ITS software layer to interoperate with the knowledge one. Thus, the software layer can be built and evolved independently on how the knowledge layer is constructed and evolved. So from this perspective, the Product Lines approach based on Semantic Services indeed produces some progress [16]. By adopting this approach one can obtain a high level of automation in the process of building applications in Product Lines. However, they do not address interoperability between knowledge and software. Thus the proposition of an approach based on Semantic Web technologies assumes the use of intelligent agents in the whole process of searching and sharing information. The agents are important in order to have better software adaptation and re-adaptation with knowledge being considered in various contexts [18][14].

Therefore, this approach enables an automated and dynamic evolution of the applications offered by a Software Product Line. Then, our proposal is to extend this approach to a Semantic Web technologies context in order to provide semantic annotation of all available resources in the context of an ITS, as well as the relationships between them, no matter how they were created.

Thus, in Section 1 a study of related work is described. In Section 2 the proposed architecture of the product line is presented. In Section 3 the artifacts to which the composition and evolution of ITS are performed in an automated way are semantically described.

2 Related Work

Nowadays, there are two approaches which are used for building Intelligent Tutoring Systems: Shell-Based Tolls and Frameworks. Both are not adequate to address the issues focused on this work.

Shell-based tools are developed to provide a mechanism for building ITS independently on the previous knowledge of advanced computing science concepts. This provides the authors with a building mechanism containing several built-in concepts and tools. This way, the author is demanded to know how the pedagogical strategies work, as well as the representation of the main ITS components: domain model and student model. That means the author needs to understand how each component work, so that each component can be properly configured. This approach is inefficient from the reuse point of view: each ITS has its own configuration in such a fashion their educational components are also related in a specific way. This fact compromises reuse, maintenance and evolution. The direct components reuse is impracticable given that each essential component has a particular configuration for each application instance. The maintenance is also a complex task because a change in any component implies changes in all components, hence, there are configuration and reconfiguration needs in the whole system. The evolution is one aspect that was not correctly dealt [1][2][12][3].

Frameworks are tools that have better reuse and adaptation power. It is an intrinsic characteristic in such projects and this implies that the author can reuse components in several applications. However, its necessary that: (i) The tutor author knows how to use the framework components as well as its behaves, for that system schedule is attached in the frameworks and (ii) what are the consequences
of the addition of a component in all the system. (iii) Moreover, its manipu-
lations are restrict to Software Engineers. From the maintenance point of view,
there is less effort: when a change occurs, work it is only needed in that par-
ticular system place. From the evolution point of view, the frameworks are very
restrict for all the possible functionalities that were previously defined [10][7]
[9]. However, every software layer is directly attached to the specific know-
ledge domain the system was designed. In this way, there are implementation
needs when any knowledge change occurs.

3 Our Software Product Line Architecture

Figure 1 shows the proposed architecture, including all the necessary elements
for the ITS production in large scale with simplified modeling, interoperability
and scalability based on Semantic Web technologies. Every component fol-
lows the Massayo [4] semantic models, which are the learner, pedagogical and
domain ones.

Component-Based Kernel (CBK): For each ITS implementation there is
model instances, these instances will be applied to specify exactly the ITS par-
ticularity. The instantiation of this model requires the specification of the do-
main in which the same is proposed to teach, what pedagogical strategies must
be used, what must be represented on the student model, and the specification
of the problem types that will be exposed to the learners.

Interface Component (IC): This component represents exactly how the
learner will interact with the ITS. This interface has to be designed according
to the domain specifications which it will be inserted. For example, the possible
problem variations that may be represented by students have to be implemented,
as well as all the possible behaviors the tutoring system may have when dealing
with the student;

Interface Communication Protocol (ICP): This protocol defines the in-
teraction rules between Manager and Interface. In fact, it is possible to have
several Interface implementations, on the other hand, in order to each interface
implementation communicate correctly with the Manager, it is necessarily for
the interface to properly implement the protocol. This protocol, besides of de-
fining the rules that establish when the interface may suffer changes from the
system (originated from the Manager), and when interface may suffer change
by the learner. It will also specify behaviors semantic and rules, in order that the
own interface can be embedded with some intelligence level;

Pedagogical Multi-Agent System: The pedagogical agents are autonomous
entities which have a collective purpose of teaching the best possible way the
learners in the system. Another important point is that each pedagogical agent
has as function a single activity well defined. Beyond this, a pedagogical agent
has the power of adaptativity related to the ontologies specification which are stored on the repository. However, it is possible to exist differences between the agents to decide which approach might be best adopted in each moment of interaction with the student. As several different pedagogical strategies may be defined in a product line specification, there may exist a negotiation between these agents to define the correct action of the system at each moment.

Pedagogical Evolution, Adaptability and Variability Protocol (PEAVP): This protocol defines the interaction rules between Pedagogical Agents. With this protocol, an agent will know the right moment to realize the environment, to act, and to invoke an other agent to help in the accomplishment of its activity. Beyond this, that protocol will be capable to define the actuation rules of the agent when there is the possibility of competition between more than one agent in choosing what is the right action to be taken. As quoted before, there may exist disagreement on choosing the best action to be taken by the system.

4 Semantic Artifact Specification and Relationships

In this section, the artifacts related to the proposed Product Line will be discussed. These artifacts are semantically described in the Ontology Descriptions, presented in Figures 1 and 2. Through these codes, all the possible educational resources will be semantically annotated. This is important in order to exist an automated process of building and instantiation of the Intelligent Tutoring Systems. Those artifacts compose the Component-Based Kernel (CBK), and they are described below:

IntelligentEducationalEnvironment: This component represents the environment as a whole. Students, teachers and tutors see this environment as a single entity, and they interact directly through it transparently.

PedagogicalIntelligentAgent: This component represents a type of agent responsible for giving instruction in a particular knowledge domain to a certain student. This agent has its own teaching strategy, it performs the planning of the learner study and monitor his activities during the process of teaching him.

SupportAgent: This component represents the agents which will support the effective work of the Pedagogical Agents. Actually, this agent will be responsible for maintaining the environment with external information to the interaction between the participants and the environment.

EducacionalResource: This component represents all the educational resources that could be used with the purpose of teaching. All resources should be linked to a specification so that they can be found and used properly and orderly.

StudentModel: This component represents the student model, it is used to identify the characteristics of each student. And through this model, the system resources will personalize the student results.

KnowledgeAssessmentMethod: This component represents the methods that assess at which cognitive level, each student who is interacting with the environment, is. Therefore, for each round of interaction between the student and the environment must exist a method that assesses the knowledge evolution of the students. These methods may follow different theories, such as constructivism, evolutionary and subjectivism.
**InferenceMechanism:** This component represents the inference mechanisms that are used in order to measure what are the possible approaches that students should receive prior to the execution of them. Aiming to optimize the education of the students, causing him to reach your goals as quickly as possible.

**DomainModel:** This component represents the syntax and semantics which a particular domain is represented. It represents all the aspects that the student needs to learn until the end of all their interactions.

**TeachingPlan:** This component represents a resource that describes at which sequence the students must interact with each educational resource. Therefore, depending on the situation in which the student is, the teaching plan should be adjusted to give continuity in the teaching process properly.

**PedagogicalStrategy:** This component represents the pedagogical strategies in which a particular pedagogical agent will compromise to address a certain student throughout the teaching process.

| 01 : SupportAgent ⊑ IntelligentAgent |
| 02 : PedagogicalIntelligentAgent ⊑ IntelligentAgent |
| 03 : PedagogicalIntelligentAgent ⊑ EducacionalResource |
| 04 : AudioClass ⊑ EducacionalResource |
| 05 : ModalLogicBased ⊑ StudentModel |
| 06 : PropositionalLogicBased ⊑ StudentModel |
| 07 : OntologyBased ⊑ StudentModel |
| 08 : Constructivist ⊑ KnowledgeAssessmentMethod |
| 09 : AnalogyStrategyBased ⊑ PedagogicalStrategy |
| 10 : SolvingProblemStrategyBased ⊑ PedagogicalStrategy |
| 11 : KnowledgeAssessmentMethod ⊑ T |
| 12 : StochasticMethod ⊑ InferenceMechanism |
| 13 : LogicalModelBased ⊑ DomainModel |
| 14 : OntologyModelBased ⊑ DomainModel |
| 15 : NeuralNetworkModelBased ⊑ DomainModel |
| 16 : BayesianNetworkModelBased ⊑ DomainModel |
| 17 : Objetivistic ⊑ KnowledgeAssessmentMethod |
| 18 : Rule-BasedReasoning ⊑ InferenceMechanism |
| 19 : Case-BasedReasoning ⊑ InferenceMechanism |
| 20 : Subjetivista ⊑ KnowledgeAssessmentMethod |
| 21 : VideoClass ⊑ EducacionalResource |
| 22 : SupportWebService ⊑ WebService |
| 23 : PedagogicalWebService ⊑ EducacionalResource |
| 24 : PedagogicalWebService ⊑ WebService |

**Ontology Description 1:** Inheritance relationship among all the entities.

Furthermore, in order to represent the way these elements relate itself, a hierarchy of entities using the concept of inheritance was created. For example, in Ontology Description 1, the relations 02 and 03 express that an Intelligent Pedagogical Agent (IntelligentPedagogicalAgent) can be seen as an intelligent agent or as an educational resource (EducationalResource). This definition is important because in the process of the construction of the Intelligent Pedagogical Agent for the environment, it should behave both as an intelligent agent (respecting interaction protocols) as educational resource (Subject to a certain
meta-definition on how to specify this type of element). The other relations of the Ontology Description 1 below comply with the same line of reasoning.

However, it is also necessary to relate the elements Computational involved in the construction of educational environments adaptive and semantic level of composition, because they define where they should be embedded elements ranging from adapting to the personalized education plan. The Ontology Description 2 represents the relationships between the entities at that level. The Following is the meaning of the described compositions relations:

**environmentHasSupportAgent:** This relationship indicates that a Intelligent Educational Environment \((\text{IntelligentEducationalEnvironment})\) has Agents Support \((\text{SupportAgent})\), according to the relations 33 and 34. In relationship 33 is expressed as the element \(\text{environmentHasSupportAgents}\) is a subset of the entity \(\text{IntelligentEducationalEnvironment}\), and the same time, in relation 34, it is stated that the same element \(\text{environmentHasSupportAgent}\) has linked to it from the elements type \(\text{SupportAgent}\). Therefore, the environments must have agents support in order to assist teaching activities promoted the pedagogical agents, such as seeking information external environment, educational resources index, updated domain model, update the model student and so on.

**environmentHasIntelligentPedagogicalAgents:** This relationship indicates that an Intelligent Educational Environment \((\text{IntelligentPedagogicalEnvironment})\) has Agents Teaching \((\text{IntelligentPedagogicalAgents})\) as the relations 35 and 36. 35 is expressed in relation to the element \(\text{environmentHasIntelligentPedagogicalAgents}\) is a subset Element Intelligent Educational Environment \((\text{IntelligentPedagogicalEnvironment})\), and that even in about 36, may have elements of type \(\text{IntelligentPedagogicalAgents}\). The idea is that the environment has bound to it agents capable of Education of policy knowledge that he intends to teach. In reality, agents Pedagogical \(\text{IntelligentPedagogicalAgents}\) built in a flexible manner in order to adjust to a given field of knowledge and a model student.

**environmentHasDomainModel:** This relationship indicates that an Intelligent Educational Environment \((\text{IntelligentEducationalEnvironment})\) has only one domain model \((\text{DomainModel})\) according to the relationships 37, 38 and 39. In the relationship 37 is described that there is only one element of type \(\text{environmentHasDomainModel}\), and this element is a subset of the entity Intelligent Educational Environment \((\text{IntelligentEducationalEnvironment})\) according to relation 38. And in relationship 39, is described that an element of type \(\text{environmentHasDomainModel}\) should have linked to it an element of the Domain Model type \(\text{DomainModel}\). This relationship is important to restrict an educational environment to just one knowledge domain. Indeed, this restriction will be chained to the process of building educational environments, because as the environment itself is restricted to a single domain knowledge, all other elements will follow this principle.

**environmentHasStudentModel:** This relationship indicates that an Intelligent Educational Environment \((\text{IntelligentEducationalEnvironment})\) has a single student model \((\text{StudentModel})\), according to the relationships 40, 41 and 42. In the relationship 40 is expressed that there is a single element of the type \(\text{environmentHasStudentModel}\), and in its turn, in the relationship 41, it represents that the same element is a subset of the element \(\text{IntelligentEducationalEnvironment}\). And in the relationship 42, it is expressed that this element has a Student Model \(\text{StudentModel}\) linked to it. In the same way that the domain model, the
specified student model for the domain will chain in a sequence of adjusts in all the elements that use them as reference to perform their tasks.

**pedagogicalAgentsHasPedagogicalStrategies**: This is the ratio representing the Pedagogical Agents (IntelligentPedagogicalAgents) must have strategies teaching (PedagogicalStrategies). This entity is defined in relations 25 and 26. 25 is represented in relation to there is a link called (pedagogicalAgentsHasPedagogicalStrategies) that is built within the Intelligent Agent for Pedagogical (IntelligentPedagogicalAgents). In turn, the relation 26, that every relationship is described has pedagogicalAgentsHasPedagogicalStrategies pedagogical strategy. This educational strategy must be implemented the following model specification of the user model user, and educational resource, so that it can be adapted to the knowledge being entered by their Agent.

**pedagogicalAgentsHasInferenceMechanism**: This is the relationship which represents that the Pedagogical Agents (IntelligentPedagogicalAgents) must have inference mechanisms (InferenceMechanism). This entity is defined in the relationships 27 and 28. Following the same reasoning to other compositions, these relationships are represented that an entity (pedagogicalAgentsHasInferenceMechanism) is a subset of Intelligent Pedagogical Agent, and that it has a mechanism inference associated with the purpose of making assumptions about the best next approach that must be accomplished to the student for each interaction. In the same way as the other entities, the actions taken by an element of this type must be customized with respect to the domain model and the student model.

**pedagogicalAgentsHasAssessmentMethod**: This is the relationship which represents that the Pedagogical Agents (IntelligentPedagogicalAgents) must have assessment methods (AssessmentMethod). This entity is defined in relationships 29 and 30. Also following the same reasoning to the other compositions, in these relationships are represented that an entity (pedagogicalAgentsHasAssessmentMethod) is a subset of Intelligent Pedagogical Agent, and that it has a assessment method (AssessmentMethod). Still, this assessment method will have as responsibility to evaluate the students’ development in relation to the interactions that are occurring between the student and the environment.

**pedagogicalAgentsHasTeachingPlan**: This is the relationship which represents that the Pedagogical Agents (IntelligentPedagogicalAgents) must have teaching plans (TeachingPlan). This entity is defined in relationships 31 and 32, which means (by analogy to other relationships) that an element pedagogicalAgentsHasLearningPlan is embedded within the element IntelligentPedagogicalAgents, and that it has the element of the type TeachingPlan. The Plan of teaching represents how the agent will lead the teaching to the student who the same will be dealing with. In it, all the resources that will be used to teach the student are related, and in which sequence of these resources will be available.

**TeachingPlanHasEducationalResource**: This relationship represents that the Teaching Plans (TeachingPlan) should have Educational Resources (EducationalResource). This entity is defined in relationships 43 and 44, which means (similarly to the other relationships) that an element TeachingPlanHasEducationalResource is embedded within the element TeachingPlan, and that the same element have the type EducationalResource. As previously mentioned, the teaching plan is who references the educational resources, and all educational resources are accessed through what is described in the teaching plan.
5 Conclusions and Future Works

This work proposes a new approach to building Intelligent Tutoring Systems in large scale concerning the knowledge as the system’s main artifact, thus, making the building process dependent on how this knowledge was conceived and distributed. After a deep study about the whole building process of the current approaches and all the resources which can be used on its building, a semantic description to enable the automatic building and evolution of ITS was created, as described by section 4. This description is part of a wider architecture which comprises the software product line architecture proposed on section 1. This architecture manifests the idea of protocols to enable the evolution and interaction through agents.

This new approach to construct ITS environments benefits to the author, in the sense that programming background is not necessary, could be verified in this proposal. From the implementation point of view, one could notice the software product line approach is, nowadays, the largest considering the software artifacts reusability, which is promising to the intelligent tutoring systems context. As future work, a refinement of the specification through the addition of new kinds of resources along with their corresponding implementations, such as pedagogical strategies and new visualization forms.

References