Authoring and delivering adaptable learning objects in SINTEC

Valentin CRISTEA, Stefan TRAUSAN-MATU
Politehnica University of Bucharest
Computer Science and Engineering Department
313, Splaiul Independentei
Bucharest, Romania

Abstract. The paper presents the solutions adopted in SINTEC, a knowledge-based e-Learning framework, to support authoring and delivering adaptable learning content. One of the main features is personalized training using enterprise / institutional knowledge repositories developed with knowledge management tools based on Web services and XML technologies. SINTEC includes tools for content development and packaging of learning objects, which can be dynamically adapted to student profiles tracked in the delivery process. The framework is compliant with different learning models, such as self-study, collaborative learning, just in time learning, and learning on demand. The environment uses e-Learning standards, such as IMS and SCORM and is based on open Web application technologies.

Introduction

The paper presents the concepts and ideas behind the authoring and delivery of adaptable e-learning content. The authors focus on content authoring and learning personalization, and on how they can be achieved by using semantic Web technologies. These concepts have been applied in the design of the SINTEC framework, which is under development at the National Center for Information Technology in the University Polithenica of Bucharest. The framework incorporates software tools for: content creation and reuse from the Internet, intelligent search of learning materials on the Web, knowledge extraction and summarization, and intelligent tutoring. They are based on open standards and technologies (including the IMS standard for e-learning). The environment may be used in various training scenarios, ranging from simple support of courses and lectures, to virtual classes and complex intelligent tutoring processes.

Developing adaptive personalized eLearning has been approached in several works. In [1] the authors describe the Adaptive Course Construction Toolkit (ACCT), which offers support based on pedagogy, instructional design principles, knowledge domain ontology descriptions and learning resource selection. In [2] the authors present the ongoing research to personalize the learning experience through adaptive educational hypermedia.

Intelligent Tutoring Systems are using students' models that typically include students' knowledge ([8]). Such models are constructed in relation with a declarative knowledge base of the considered domain, which is acquired through a conceptualization process, the result being an ontology. Such ontologies may be reused for many different applications and they are easily extensible. Several standards exist related to Web languages for ontology exchange (e.g. OWL) or languages for representing such data (DAML, DAML+OIL).

Our approach is combining the ideas of ITS, Cognitive Psychology, dynamic generation of Web pages ([7]), Emotional Intelligence (EI, Goleman) with the facilities offered by reusable ontologies on the Web and rule-based programming (e.g. CLIPS). The tools for content creation and dynamic delivery adaptation were developed as components of the SINTEC framework, but their re-design as as open Web services permits their shipping with the learning objects and easy integration with client's platform.

The structure of the paper is as follows. Section 1 presents the architecture and components of the SINTEC environment. It serves as a basis for understanding the description in Section 2 that refers to the tools for content creation and dynamic delivery adaptation. Section 3 presents conclusions and future work.

1. SINTEC architecture

Personalization is a key premise for an improved learning experience. Personalization is closely associated with e-learning and refers to the following issues:

- **Interface personalization** can range from presenting some items on user's display in accordance with user's options, up to more complex processes that include establishing the user emotional profile and adapting the interface according to the result.
- **Content personalisation** involves authoring adaptable learning materials, constant evaluation of student's knowledge level, and adaptation of learning materials accordingly. Evaluation may not always be proactive from the user's perspective (such as on-line assessments). Advanced techniques have been studied such as text mining applied to the user's public messages.

Due to the use of knowledge-based technology that includes not only domain knowledge but also psychological and pedagogical knowledge, SINTEC is able to intelligently tailor the learning process according to users' profile and progress. This makes it very useful for companies and institutions where the users profile and knowledge vary in a very large range. Moreover, it can be used around any particular domain ontology.

The framework includes a collection of tools and repositories that integrate collaborative techniques on the web with knowledge-based methods, and multiple purpose XML-based annotation (metadata, exchange of reusable components, knowledge representation) that empowers personalization. SINTEC dynamically builds and monitor ontology-based learner models, which can be further used to adapt the instruction strategies (sequences of learning objects) to learner characteristics. In addition, it tracks students' activities and interactions with the learning material, analyses their answers and texts they write, identifies needs or interest and evaluates students' psychological profile and learning style. Socio-emotional intelligence issues are also considered for tailoring the learning process [3].

Content creation and reuse is one of the main problems for e-learning systems. Professors find difficult to develop learning modules according to e-learning standards. Particular problems are the structuring and the organization of learning materials in conceptual units and the addition of metadata definitions. Other requirements for e-learning are:

- The addition of new knowledge to the knowledge pool should be easy and straightforward to persons not familiar with advanced IT technologies. Therefore, the architecture requires two different types of ontologies:
    - A *pedagogical ontology* containing concepts such as learner, learning task, activity, grade, etc. This contains knowledge about the structure and usage of the knowledge system itself.

    - Several *domain ontologies*, each containing the knowledge pool for a certain domain, such as Algorithms Analysis, Compilers, Operating Systems, etc.
- The development of new domain ontologies should be straightforward even for teachers without a technical background. This is quite difficult to achieve and the viable method will be not only sets of clear instructions and samples about how to build an ontology, but also specialized editors.

The SINTEC platform is currently applying the idea of the above approaches; therefore it allows the exchange of the following types of information with other similar applications:

- Exchange of the user profile and background information, including estimated knowledge level and full training history is accomplished through the implementation of the IMS Learner Profile Information (LIP) specification
- Exchange of several types of learning content (e.g. lecture notes, practical exercises, course support materials) is accomplished through the use of IMS Metadata and Content Packaging specifications
- Exchange of test information, including questions, tests, grading and evaluation information, as well as full result history is accomplished through the use of IMS Question & Test Interoperability (QTI) specification.

The architecture of the SINTEC system is illustrated in Figure 1. From the knowledge perspective, it comprises three main groups of modules:

- Content creation and management,
- Knowledge server
- Intelligent tutoring

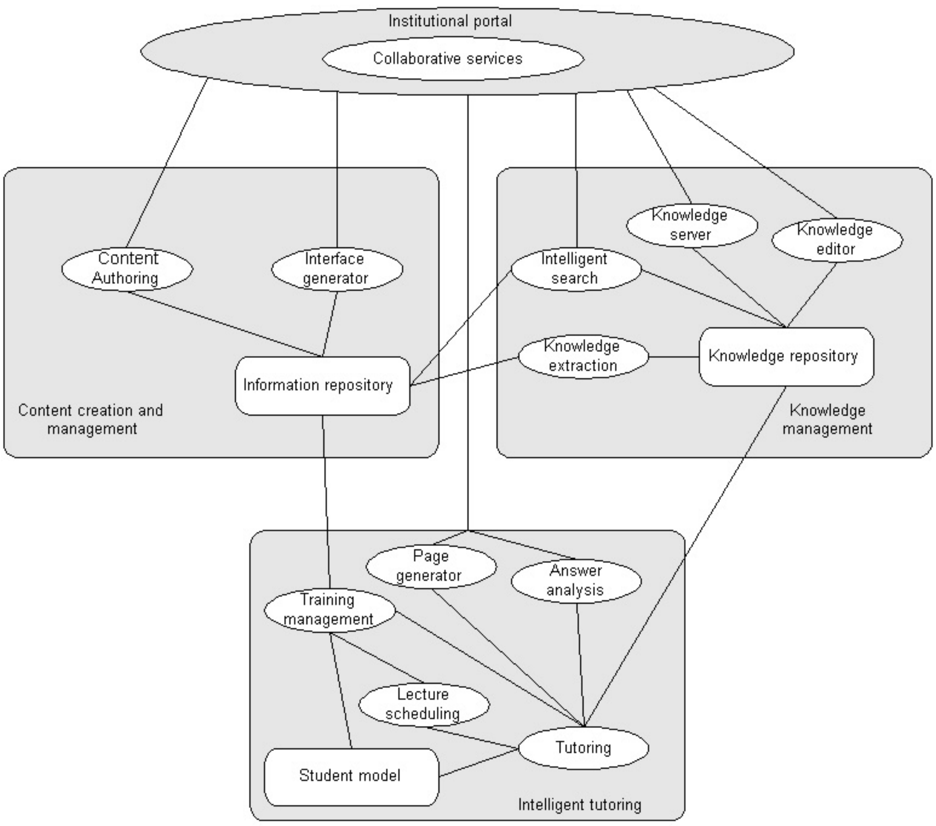

Figure 1. SINTEC architecture

For providing knowledge-based, intelligent tutoring, adaptive learning tools for the development of student models are used [3]. These tools track student's activity and her interactions with learning materials, analyze her answers and texts she writes, identifies needs or interest and evaluate her psychological profile and learning style. One important component of the student model is the knowledge level: what knowledge she has, what knowledge she does not have and what knowledge she has wrongly. These facts are derived from answers to different questions, from the analysis of essays written by students, from students' interactions. The facts may be further used for dynamic web pages and test generation and lesson planning.

SINTEC uses several types of ontologies. The most important is the *domain* ontology that includes the basic concepts and relationships in the domain taken into account. In addition, an ontology for *pedagogy* is used for the generation of flexible, personalized learning processes. This ontology includes also the Bloom taxonomy.

A part of a *cognitive psychology ontology* that contains the concepts related to learner profiles and emotions is used for assuring a user-friendly human-computer interaction. For the selection of the relevant documents in a given learning context and of the manner of presentation, a *document ontology* is useful. Such an ontology is helpful also for the processes related to handling documents repositories and text mining.

In addition to declarative, ontological knowledge, procedural knowledge is represented in SINTEC by using production rules (in a Jess-like language offered by the Protégé environment). These rules refer to concepts, relationships and individuals in the above ontologies. They use data about the learner (e.g. his actions or results at tests) to infer and update the learner model and for planning the next learning steps.

For example, such a rule might say that:

IF the learner has not obtained a given score at a given test
AND the learning process has not been longer than the possible limit of learner's interest and attention
THEN give to the learner the task of reading the modules (or some dynamically generated web pages [6]) that covers the concepts that resulted (by inference using specific rules) to be not understood from the tests.

The training environment:

- provides a flexible and easy to use platform for both students and tutors;
- uses adaptive content based on user preferences and preparation level, both for course and test preparation and analysis;
- adapts easily to a specific domain by incorporating an adequate specific ontology;
- provides interoperability with other applications conforming to a similar set of standards contain both presentation and content authoring services Flexible, standardized, adapted to enterprise needs and to trainees profiles (including emotional intelligence).

The design aimed at obtaining platform independent components that permits a rapid deployment on different premises.

2. Tools and technologies for authoring and adaptation of learning content

This section will discuss the tools that are to be shipped with the learning objects and some of the technologies used to create learning objects and the tools themselves. We selected the virtual class scenario, in which a tutor prepares learning materials for a class and then adapts them to students according to their profiles. A tutor has the necessary tools to:

- Repackage the learning objects in the organization form best suited for the trainees.

- Define the sequencing information for the new organization.
- Alter the original metadata for revisions, authors, technical requirements, student prerequisites, etc.
- Track student progress through the activities of a course and programmatically obtaining the sequence of activities best suited for a student based on the sequencing information associated with a course package.

In order to simplify the applicability of these functions, the learning objects are shipped with two different applications:

- The *Content Authoring Center* that deals with the first three issues related to content and packaging of learning objects.
- The *Tutor Aid Center* which constitutes a tool for the instructor at delivery-time (i.e. during the course).

Figure 2 clarifies this organization. The tutor aid helps to interpret sequencing information that is not in a readable form. The IMS Simple Sequencing specification is quite complex and contains several types of rules:

- *Rollup-rules* specify how the result of a **child activity** (which is part of a larger activity) is reflected in a result for the **parent activity** (some of the simplest possible rules are: *Satisfied-If-All-Satisfied* - the parent activity is completed if all parts are completed; *Satisfied-If-Any-Satisfied* - the parent activity is completed if any part is completed).
- *Selection rules* specify how future activities are selected based on previously collected progress information. This, in turn, leads to the possibility of personalized content delivery.
- *Sequencing rules* specify how the selected activities are ordered and how the trainee may undertake them (in sequential order, by choice, etc.)

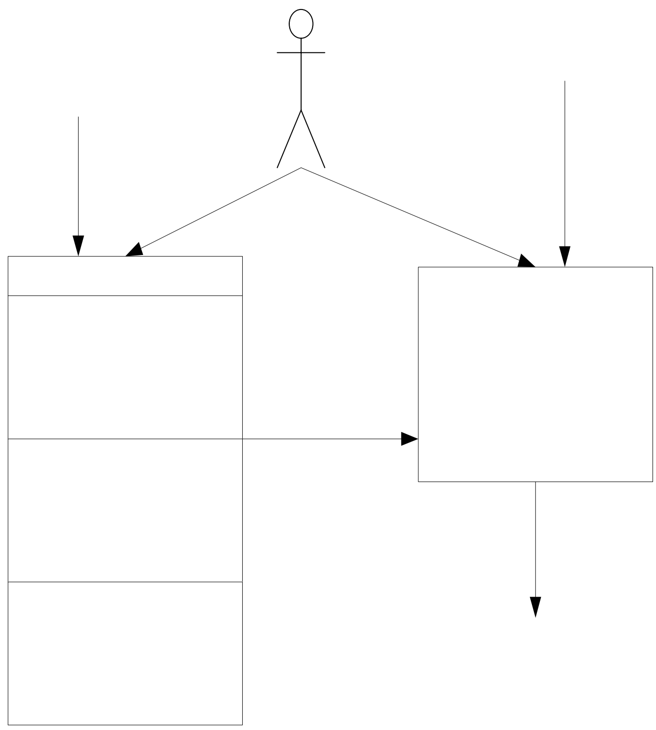

Figure 2: Instructor tools

Therefore, a tutor may find difficult applying and computing all this information manually. The *Tutor aid* helps an instructor to programmatically select the best activity track for each student, based on the information on student progress and the sequencing information found in the learning objects package. On the other hand, the *Content Authoring Center* helps defining a package structure from the learning objects provided. This includes the basic organization, the metadata information and the sequencing information for the three standards previously discussed.

SINTEC makes use of the IMS metadata specification to represent knowledge about an item. The format chosen for the knowledge representation for a learning item (for example, a web page or a presentation document) is a string representation of a hash table, in <Ci, ri> pairs. **Ci** is the actual name of the concept addressed by this piece of learning material, while **ri** is a percentage describing relevance of that concept in the context of the current material. Since the concepts themselves are part of the domain ontology, the metadata editor is provided where instructors can simply choose the desired concepts from a knowledge base. In the simplest case, the relevance is automatically considered to be equal for all concepts involved, but the instructor may use different other policies. In Figure 3, the approach taken in SINTEC for relevance dynamic estimation is presented.

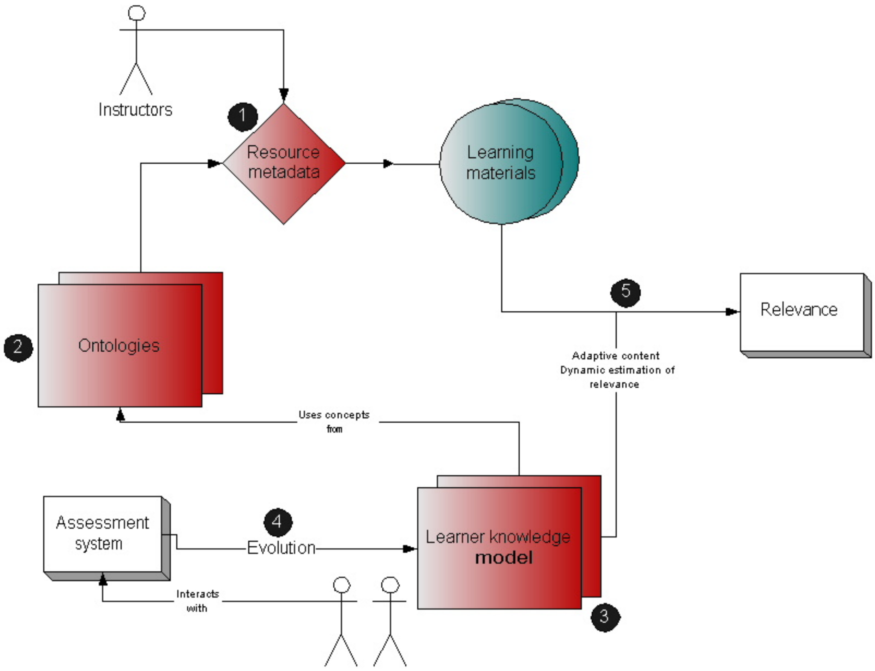

Figure 3. Dynamic estimation of relevance

The model used in the *Tutor aid* for representing user's knowledge is based on the notions of learning activity and action. A *learning activity* is an interaction between the learner and the system, which has a clearly defined goal of either transferring or assessing the transfer of a unit of knowledge to the learner. Examples are: reading learning materials, homework, group projects, taking an online assessment. An *action* is a piece of the learning activity that produces a feedback to the system that can be quantified. An example of such an action is answering a question in an online assessment.

We also define a *state* for each concept in the user knowledge data, as well as a positive and negative *score*. The states may include: unconscious-known, wrongly-known, self-learner, well-known, etc. The scores come from the quantification of learning activity actions. The number of actions providing positive or negative score is also necessary for computing averages. An example of a user concept network is depicted in Figure 4.

The network can “evolve” upon the occurrence of new positive or negative actions. When new scores arrive, they are added to the concepts addressed by the action; at least some portion of the new score must be propagated to the neighboring concepts. New concepts may be added if they are in a direct relationship to the concept whose score increases. The *propagation algorithm* we use always guarantees a limit to the number of propagation steps. This is achieved through the gradual diminishing of the propagated value until it falls under a constant threshold.

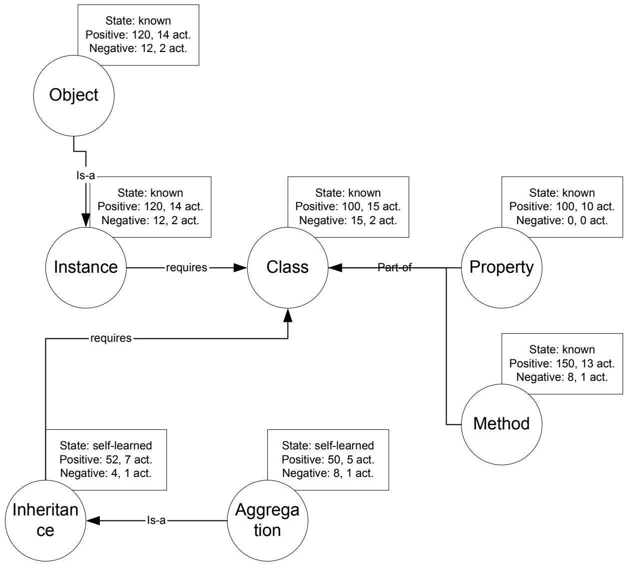

Figure 4. Part of a concept network

The querying and updating of the concept networks and ontology system is done through Algernon, a *rule-based system* developed at the University of Texas at Austin. This system has a good Java API and interface very well with the Protégé ontology engine. Since the updating of the knowledge information is done asynchronously, the production rules can be used to govern these processes.

The *Tutor aid* can use the knowledge information in several ways. Two are mentioned here: providing a personal agent for each learner that gives recommendations on certain learning materials and/or activities that a student can use to improve his/her overall knowledge level; providing a *personalized activity tree* for each learner (solution adopted in SINTEC). For example, an individual with some advance knowledge may skip some activities that he/she already covered (as indicated by his/her knowledge level).

One important issue pertaining to the proposed implementation is the means of actually developing these tools and a discussion of the technologies involved. First of all, the tools shipped together with the learning objects rely heavily on XML-based technologies (since the IMS specifications themselves are based on XML). We have chosen Java as the implementation language because of its inherent portability and its applicability to Web-based applications. Furthermore, we recognize that the intended target for this environment (that is, other e-learning research and development centers, in the academia or elsewhere) already have in place some software and hardware platforms intended for training purposes. There are two important issues here:

- Since the tools are developed in Java and use XML, the problems pertaining to compatibility are greatly reduced.
- However, with the goal of interoperability in mind, we do not impose the use of specific applications to the mentioned audience. Therefore, another version of the same tools described here *will be available as Web Services*. This greatly increases the possibility of integration with the client platforms. Besides the standalone Java versions, the URL endpoints of these Web Services that provide the same functionality, as well as the development information and Java clients for the named Web Services will be shipped as part of the same package.

3. Conclusions

As mentioned in the introduction, the experience with online course delivery of the National Center for Information Technology has led to several stages of development for an e-learning environment intended for both students and the general public. The proposal herein is the latest development from this center and focuses on advanced learning techniques and technologies. This environment attempts to achieve greater quality of learning through:

- High quality of learning materials;
- Personalized content delivery;
- Adaptation of course structure and additional information to a variety of scenarios.

Some of the most important goals in the e-learning world are reached and implemented by this environment, namely:

- Flexibility – the curriculum can adapt both to different scenarios and to a great number of student profiles within one given scenario.
- Extensibility & modularity are directly derived from the learning object structure and the ease of re-packaging these objects into flexible structures.
- Interoperability is achieved through the use of open standards such as IMS and XML, as well as the use of Java and Web Services for the implementation of tools.

References

[1] Dagger, D., Wade, V., Conlan O. (2004), *A Framework for Developing Adaptive Personalized eLearning,* retrieved May 2005 from https://www.cs.tcd.ie/Owen.Conlan/publications/elearn2004_daggerd.pdf

[2] De Bra, P., Aroyo, L., Cristea, A. (2004), *Adaptive Web-Based Educational Hypermedia*, retrieved May 2005 from http://www.win.tue.nl/~debra/webdynamicsbook/paper.pdf...

[3] Conati, C., Zhou, X., *Modeling students' emotions from cognitive appraisal in electronic games*, in Cerri, S., Gouarderes, G., Paraguacu, F. (eds.), Intelligent Tutoring Systems, Springer 2002, pp.944-954.

[4] Dimitrova, V., Self, J., Brna. 2000. *Maintaining a Joinly Constrcted Student Model.* In S.A.Cerri (ed.), *Artificial Intelligence, Methodology, Systems, Applications 2000*, Springer-Verlag, ISBN 3-540-41044-9, pp.221-231.

[5] Sowa, J. (1999), *Knowledge Representation: Logical, Philosophical and Computational Foundations*, Brooke Cole Publishing Co., Pacific Grove, CA,.

[6] Trausan-Matu, St., Maraschi, D. and Cerri, St. (2002), *Ontology-Centered Personalized Presentation of Knowledge Extracted From the Web*, in S.Cerri, G.Gouarderes (eds.), Intelligent Tutoring Systems 2002, Springer, Lecture Notes in Computer Science number 2363, pp 259-269.

[7] *Question & Test Interoperability Specification v 1.2.1*, IMS Consortium, 2003

[8] G.H. Von Wright (1971), *Explanation and Understanding*, London: Routledge & Kegan Paul

[9] *Content Packiging Specification v 1.1*, IMS Consortium 2003

[10] *Metadata Specification v 1.2*, IMS Consortium 2003

[11] *SCORM Final Specification v2.0*, Advanced Distributed Learning Consortium

[12] Cerri, S., Gouarderes, G., Paraguacu. 2002. *Intelligent Tutoring Systems.* Springer.

[13] Sleeman, D., Brown, J.S. 1982. *Intelligent Tutoring Systems*, Academic Press, 1982.

[14] P. Ciancarini et al., Coordinating Multiagent Applications on the WWW: A Reference Architecture, *IEEE Trans on Software Engineering, 24* (5), May 1998, 362-375.

[15] G.H. Von Wright, *Explanation and Understanding* (London: Routledge & Kegan Paul, 1971)

[16] Chepegin,V., Aroyo, L., De Bra, P., Heckmann, D. (2004), *User Modeling for Modular Adaptive Hypermedia*, retrieved May 2005 from http://wwwis.win.tue.nl/~vchepegi/publications/SWEL04.pdf