Semantic Web based Adaptive Intelligent Tutoring System

Suraiya Jabin¹, K. Mustafa²

¹,² Department of Computer Science, Jamia Millia Islamia, Central University, New Delhi, India
¹Email: sjabin@jmi.ac.in
²Email: kmfarooki@yahoo.com

Abstract—In the last several years, Web-based education has become a very important branch of educational technology. Classroom independence and platform independence of Web-based education, availability of authoring tools for developing Web-based courseware, cheap and efficient storage and distribution of course materials, hyperlinks to suggested readings, digital libraries, and other sources of references relevant for the course are but a few of a number of clear advantages of Web-based education. However, there are several challenges in improving Web-based education, such as providing for more adaptivity and intelligence. The main idea is to incorporate semantic web tools and resources to the design of artificial intelligence in education (AIED) systems aiming to update their architectures to provide more adaptability, robustness and richer learning environments. However, the construction of such systems is highly complex and faces several challenges in terms of software engineering and artificial intelligence aspects. This paper presents state of the art semantic web methods and tools used for modeling and designing intelligent tutoring systems (ITS). Also it tries to draw attention of Semantic Web users towards e-learning systems with the hope that the use of semantic web technologies in educational systems can help the accomplishment of Anytime, Anywhere, Anybody Learning where most of the WWW resources are reusable learning objects supported by standard technologies and learning is facilitated by intelligent pedagogical agents.

Index Terms — AIED, ITS, Learning Objects, Learning Management Systems, Ontology, Semantic Web.

I. INTRODUCTION

Intelligent tutoring systems (ITS) appeared during the 1970s, driven by the success of knowledge-based systems and expert systems. ITS are able to instruct and train students and professionals without the intervention of human beings. ITS introduced a set of ideas, like the use of computational models of domains, allowing the possibility of reasoning and explaining domain problems automatically. Developments were made in trainees’ models, instructional and pedagogical planning, and user interface. In the 1990s, with the Web boom, some ITS ideas were incorporated in new computer-aided instruction paradigms, like e-learning and distributed learning. However, there is a clear difference in the level of interactions and types of skills addressed by ITS and other e-learning systems. The use of artificial intelligence techniques to educational software design influenced the evolution from Computer Assisted Instruction (CAI) to Intelligent Tutoring Systems (ITS) or Intelligent Computer Assisted Instruction (ICAI). Computer-based Instruction (CBI) and Web-based Instruction (WBI) are the two primary instructional or ‘courseware’ products with the understanding that CBI utilizes CDROM (or non-Web-based) technology to deliver its courseware, and WBI utilizes Internet (or Web-based) technologies. The type of delivery medium has played an important role in determining what instructional designs are possible.

The system consists of several core modules, which are relatively independent of each other to allow easy upgradability and portability to other teaching domains [2]:

1. Personalized Student/Learner Model, which monitors the progress of every individual student. Learner modeling module is composed of mechanisms that acquire and represent the learner’s knowledge about a specific subject domain.

2. An Expert/Domain Knowledge Model, basically contains a domain knowledge base and some mechanisms to reason about this knowledge. Generally, this module is responsible for problem solving tasks, using some resources from AI, like logic, production rules, semantic network, frames and bayesian networks.

3. Pedagogical/Teacher module is responsible for selecting resources from a domain as well as deciding about the pedagogical action to be accomplished during the interaction process with the learner.

Communication /Interface module is responsible for directly managing the interactions with the learners.

One of the hottest R&D topics in recent years in the AI community, as well as in the Internet community, is the Semantic Web. The World Wide Web is a collection of electronic documents linked together like a spider web. These documents are stored on servers located around the world. Web contains virtually boundless information in the form of
documents. The semantic web (SW) extends the classical web in the sense that it allows a semantic structure of web pages, giving support to humans as well as artificial agents to understand the content inside the web applications. As a result, Semantic Web provides an environment that allows software agents to navigate through web documents and execute sophisticated tasks.

SW itself offers numerous improvements in the context of Web-based educational systems contributing to the upgrade of learning quality. Tim Berners-Lee [1], inventor of the Web, proposed the term Semantic Web, and defined the Semantic Web as follows: “The Semantic Web is not a separate Web but extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. This is a web of data that can be processed by machines directly and indirectly”.

This paper presents a survey of different architectures used to combine Semantic Web techniques with ITSs, explaining its elements, relationships, challenges, and the different design criterions, offering some guidelines to make decisions when different implementation solutions are possible.

The remainder of this paper is organized as follows: Section II presents different successful semantic web methods and tools used for modeling and designing intelligent tutoring systems (ITS), section II discusses the adaptivity issue in Web-based Educational System, section III presents recent developments in the area, Section IV gives comparative study and analysis of different methods discussed, Section V presents conclusions and future work.

II. ITS: A SEMANTIC WEB PERSPECTIVE

Semantic Web perspective tries to incorporate new era of intelligent learning infrastructures in knowledge intensive organizations for example: knowledge management, learning management, e-collaboration, ontologies, semantic web technologies, corporate learning, integrative knowledge portals, distributed knowledge and learning repositories, peer-to-peer technologies etc.

The design and implementation of Intelligent Tutoring Systems (ITS) is a very complex task, as it involves a variety of organizational, administrative, instructional and technological components. In addition, there are no well established methodologies or development tools for ITS implementation. Therefore systematic, disciplined approaches must be devised in order to leverage the complexity of ITS implementation and achieve overall product quality within specific time and budget limits.

According to [5], the Educational Semantic Web is based on three fundamental affordances. The first is the capacity for effective information storage and retrieval. The second is the capacity for non-human autonomous agents to augment the learning and information retrieval of human beings. The third affordance is the capacity of the Internet to support, extend and expand communications capabilities of humans in multiple formats across the bounds of time and space. These fundamental affordances can be achieved through some technologies:

A. Ontologies:

information on the Web is commonly represented in natural-language for human understanding. However, in order for the computer to understand its meaning, it is necessary to represent the information in a form that can be interpreted syntactically and semantically. Such representation helps the process of analyzing, extracting, and integrating information on the Web, making it easier the creation of solid knowledge bases that intelligent services can rely on to support users’ needs. Nowadays, research on ontologies has been considered one of the keys to provide information in a computer-understandable way. In computer science, ontology represents a set of precisely defined terms about a specific domain and accepted by this domain’s community. Ontology is an explicit specification of a conceptualization. Technically, an ontology is a text-based piece of reference-knowledge, put somewhere on the Web for agents to consult it when necessary, and represented using the syntax of an ontology representation language. There are several such languages around for representing ontologies, see [3] for an overview and comparison of them. It is important to understand that most of them are built on top of XML and RDF. According to [4], the use of ontologies and the advent of intelligent services for developing Web content, Web filters, intelligent search engines, and other applications are transforming the Web of information into the Semantic Web;

B. Learning Standards:

the use of standards is fundamental in describing, developing, exchanging, accessing, annotating, combining and qualifying educational resources. Both Semantic Web standards (e.g., RDF, SKOS) and educational resources (e.g., IEEE, IMS) are used in the development of Semantic Web-based Educational System [6].

C. Semantic Web Services:

SWSs provide a number of different activities transforming a static collection of information into a distributed way on the basis of Semantic Web technology making content within the WWW machine-processable and machine-interpretable;

D. Intelligent Agents:

they are autonomous software entities which provide several kind of support through the interaction with players according to their roles. In addition, these entities must have the ability of execute and accept new requirements on the fly.

E. RDF

The RDF is a simple meta model for defining and exchanging information on the semantic web. The basis of a particular way of providing meaning for metadata is embodied in the model theory for RDF, the language at the base of the Semantic Web. The RDF Schema [1] (RDFS) draft specifies a small upper ontology on top of RDF, but it too is a work in progress and has never been officially published. Finally, an RDF Model Theory that formally defines the semantics of RDF [11] and RDFS constructs is also under development. Consequently, while the foundations of RDF
are fairly solid and well understood, the more advanced features (that nonetheless belong to the object layer of the IMI Reference Model) are still very much a moving target. In particular, RDF has a very limited collection of syntactic constructs, and these are treated in a very uniform manner in the semantics of RDF. The RDF thesis requires that no other syntactic constructs are to be used and that the uniform semantic treatment of syntactic constructs cannot be changed, only augmented.

F. XML:
The eXtensible Markup Language (XML) has become a standard language for data representation and exchange. XML [9] is a Standard, flexible syntax for data exchanging Regular, structured data. Database content of all kinds: Inventory, billing, orders etc. It has small typed values and irregular, unstructured text. It consists of documents of all kinds: Transcripts, books, legal briefs etc. With continuous growth in XML data sources, the ability to manage collections of XML documents and discover knowledge from them for decision support becomes increasingly important. Mining of XML documents significantly differs from structured data mining and text mining. XML allows the representation of semi-structured and hierarchical data containing not only the values of individual items but also the relationships between data items. Element tags and their nesting therein dictate the structure of an XML document. XML was designed to transport and store data. XML tags are not predefined. You must define your own tags XML documents [10] form a tree structure that starts at “the root” and branches to “the leaves”.

G. OWL:
By 2004, the most popular higher-level ontology representation languages were OIL (Ontology Inference Layer) and DAML+OIL [8]. An ontology developed in any such language is usually converted into an RDF/XML-like form and can be partially parsed even by common RDF/XML parsers [9]. Of course, language-specific parsers are necessary for full-scale parsing. There is a methodology for converting an ontology developed in a higher-level language into RDF or RDFS [10].

In early 2004, W3C has officially released OWL (Web Ontology Language) as W3C Recommendation for representing ontologies [7]. OWL is developed starting from description logic and DAML+OIL. The increasing popularity of OWL might lead to its widest adoption as the standard ontology representation language on the Semantic Web in the future. Essentially, OWL is a set of XML elements and attributes, with well-defined meaning, that are used to define terms and their relationships (e.g., Class, equivalentProperty, intersectionOf, unionOf, etc.). OWL elements extend the set of RDF and RDFS elements, and the owl namespace is used to denote OWL encoding. Figure 2 shows a piece of a simple ontology developed using the OWL language. In practice, ontologies are often developed using integrated, graphical, ontology-authoring tools, such as Protégé-2000, OILed, and OntoEdit [18]. They are used to develop new ontologies and modify existing ones. They let the author edit and develop ontologies concentrating on the domain’s concepts and relationships, without worrying much about ontology-representation languages. The author can choose ontologies from a list, choose attributes and relations from another list, edit, add, remove, and merge ontologies. The output is usually produced in a specific high-level ontology-representation language such as OWL, RDF/RDFS, HTML, or in plain text.

Figure 2. A simple ontology defined in OWL [11].

To support the vision of the Semantic Web which is making machine-readable content available on the Web, several software platforms and application interfaces (APIs) have been developed to permit the automatic creation and use of RDF(S) and OWL ontologies. Protégé is an open-source platform developed at Stanford Medical Informatics. It provides an internal structure called model [23] for ontologies representation and an interface for the display and manipulation of the underlying model. The Protégé model is used to represent ontology elements as classes, properties or slots, property characteristics such as facets and constraints, and instances. The Protégé graphical user interface can be used to create classes and instances, and set class properties and restrictions on property facets. Additionally, Protégé has a library of various tabs for the access, graphical visualization, and query of ontologies. Protégé can be currently used to load, edit and save ontologies in different formats including XML, RDF, UML, and OWL [22].

Intelligent Tutoring Systems (ITS) are an effective manner in order to improve the performance of students. At the moment, the most important research works on cognitive tutor are focused on the cognitive process. Nevertheless, this approach focuses on a more general way of developing ITS using learning objects and learning paths. With this in mind, it is possible to develop learning strategies for creating new learning paths through check points that measure or score the knowledge of students.

III. ADAPTIVITY ISSUE IN WEB-BASED EDUCATIONAL SYSTEM

An adaptive web-based educational system (AWBES) is a system that changes its configuration in order to improve the students learning. In other words, its goal is to provide adaptive interactions to the learners aiming to improve the quality of services [12]. The adaptation types are described as follows [13]:
- Instructional model adaptation: this adaptation form allows the student to have different content, activities and services according to the specifications made by the course author. At project time, the author may only specify which attributes a user may have, aiming to receive certain content, activity or access to a service.

© 2011 ACEEE
DOI:03.CSS.2011.2.112
At execution time, the student model has to be checked for specified conditions agreements in order to decide which content, activities and services will be provided to the student. The referred content, activities and services shall be properly modeled to apply this adaptation at execution time:

- Adaptive interactions: this adaptation form provides support to the students whilst they interact in a certain course. This support is addressed to the student and the tutor, possessing several services, contents and activities to work in the interaction. Besides, user support is given by considering the information stored in different models, especially the user, group and service models;
- Presentation adaptation: this adaptation form presents a different user interface for each student according to his or her model. This adaptation does not refer only to what the user has customized, but also to what the system has learned from previous interactions of that and other users. This is one of the most efficient ways of building the presentation for certain user learning.

IV. RECENT DEVELOPMENT IN THE AREA

Most of the Universities or Educational Institutes use VLE (Virtual Learning Environment portal) to deliver online courses. Moodle is an Open Source Course Management System (CMS), also known as a Learning Management System (LMS) or a Virtual Learning Environment (VLE). It has become very popular among educators around the world as a tool for creating online dynamic web sites for their students. To work, it needs to be installed on a web server somewhere, either on one of your own computers or one at a web hosting company. Moodle [http://moodle.org/] is one of the best VLE as it is well appreciated by educational sector. The focus of the Moodle project is always on giving educators the best tools to manage and promote learning, but there are many ways to use Moodle:

1) Moodle has features that allow it to scale to very large deployments and hundreds of thousands of students, yet it can also be used for a primary school or an education hobbyist.
2) Many institutions use it as their platform to conduct fully online courses, while some use it simply to augment face-to-face courses (known as blended learning).
3) Many of our users love to use the activity modules (such as forums, databases and wikis) to build richly collaborative communities of learning around their subject matter (in the social constructionist tradition), while others prefer to use Moodle as a way to deliver content to students (such as standard SCORM packages) and assess learning using assignments or quizzes.

SCORM is the “Shareable Content Object Reference Model”. It is a set of standard specifications that are designed to help us share learning materials between different systems. It is primarily used by, and in, Learning Management Systems (LMSes, also known as VLEs, or Virtual Learning Environments). Some well-known LMSes are Moodle, Blackboard and Sakai. The basic idea of SCORM is that, if you create a piece of learning material such as a multiple-choice exercise, you should be able to upload that material into any LMS system, and it should work: the LMS should be able to display it, and the exercise should be able to report its information (student scores etc.) back to the LMS system.

Another popular tool is Cognitive Tutor Authoring Tool (CTAT) [20] developed by the Human Computer Interaction Institute from Carnegie Mellon University. CTAT enables you to create two main types of tutors: Example-tracing tutors, which can be created without programming, but require problem-specific authoring, and Cognitive Tutors, which require programming a cognitive model of student problem solving, but support tutoring across a range of similar problems.

There are a few factors that may influence the decision regarding which tutor type to build:

- The complexity of the tutor problems (i.e., the problems that students solve as they work with the tutor);
- Whether many problems of the same type need to be developed;
- How many alternative solutions paths exist for each problem;
- Whether subtle ordering constraints exist on the steps within the tutor problems;
- Whether later steps in a problem depend on earlier steps; and
- Whether in-house expertise in AI programming is available to develop a cognitive model.

Peter Reimann, Kalina Yacef and Judy Kayin (2011) attempt to relate types of change processes that are prevalent in groups to types of models that might be employed to represent these processes. Following McGrath’s analysis of the nature of change processes in groups and teams, they distinguish between development, adaptation, group activity, and learning. Two types of event-based process analysis are discussed in more depth: the first one works with the view of a process as a sequence pattern, and the second one view a process as an even more holistic and designed structure: a discrete event model. For both cases, they provide examples for event-based computational methods that proved useful in analyzing typical CSCL log files, such as those resulting from asynchronous interactions (focus on wikis), those resulting from synchronous interactions (focus on chats) [14].

The article My Science Tutor (MyST) [23] describes, an intelligent tutoring system designed to improve science learning by students in 3rd, 4th, and 5th grades (7 to 11 years old) through conversational dialogs with a virtual science tutor. In our study, individual students engage in spoken dialogs with the virtual tutor Marni during 15 to 20 minute sessions following classroom science investigations to discuss and extend concepts embedded in the investigations. The spoken dialogs in MyST are designed to scaffold learning by presenting open-ended questions accompanied by illustrations or animations related to the classroom investigations and the science concepts being learned. The focus of the interactions is to elicit self-expression...
from students. To this end, Marni applies some of the principles of Questioning the Author, a proven approach to classroom conversations, to challenge students to think about and integrate new concepts with prior knowledge to construct enriched mental models that can be used to explain and predict scientific phenomena. This article describes how spoken dialogs using Automatic Speech Recognition (ASR) and natural language processing were developed to stimulate students’ thinking, reasoning and self explanations. It describes the MyST system architecture and Wizard of Oz procedure that was used to collect data from tutorial sessions with elementary school students. Using data collected with the procedure, we present evaluations of the ASR and semantic parsing components. A formal evaluation of learning gains resulting from system use is currently being conducted. This paper presents survey results of teachers’ and children’s impressions of MyST.

Ullrich C., Shen R., Tong R., and Tan X. (2010) presented a project that employs mobile technology to provide live access to video lectures education to the largest number of citizens possible. Motivated by the observation that in developing countries, mobile phones have a much higher penetration rate than laptop and desktop computers, they developed a mobile learning system that streams live lectures to the students’ mobile devices. The lectures are held as usual in university, not requiring the costly preparation of especially authored mobile learning materials. The system takes care of compressing the video and audio data efficiently so that it can be live-streamed, while maintaining high visual quality of the slides. They describe the system architecture and the outcome of two large-scale evaluations in two lectures with about 1,000 students at the Distant College of Shanghai Jiao Tong University (Online-SJTU), an online college with 26,000 students [15].

David Poe, Nirmala Venkatraman, Christine Hansen, and Gautam Singh (2009) presented a software Bioinformatics Foundational Learning Objects Workbench (BioFLOW) to teach Bioinformatics which is a recent, hybrid field of study combining elements of biology, statistics, and computer science. BioFLOW is a system that is intended to be used by instructors to educate their students in bioinformatics principles at a variety of grade levels. BioFLOW functions by dividing topics of bioinformatics into modules. These modules cover a wide variety of topics and may be inserted into different courses in bioinformatics-related fields, providing students with a basic understanding of bioinformatics if they choose to pursue it as a field of advanced study. Utilized together, these tools can be used to refine the content of the teaching modules and thus provide an effective system for educating students in bioinformatics [16].

Fares Fraij and Victor Winter (2008) presents an intelligent tutoring framework that can be effectively utilized to assist teaching courses and therefore to achieve pedagogical goals. The courses generated using the framework is adaptive, i.e., they adjust their behaviour to overcome the individual differences among students. The architecture of the framework provides three modules for an administrator, an instructor and a student. Furthermore, students explore the material of the course through two modes, namely non-interactive and interactive (or adaptive). To achieve the goals of the framework, it is recommended to employ an agile software development process such as extreme programming. Furthermore, the development team of the framework must involve students and therefore proceeds in a user-centered fashion [17].

IV. COMPARATIVE STUDY AND ANALYSIS

Most of the analyzed system focuses on reusing learning objects, like the systems RDF Description Model for Manipulating Learning Objects and ELENA Smart Space for Learning [21]. In general, each learning object has its own and different learning strategy that can compromise the uniformity of the system. There are systems that have a pedagogical model but do not show a clear representation of this model or the relationship among the components.

Current trends in Web technology suggest that appropriate representation languages include XML, XML Schema, RDF, and RDF Schema languages, all developed under the auspices of WWW Consortium http://www.w3.org/XML; http://www.w3.org/RDF). For developing domain ontologies, higher-level languages and graphical tools built on top of those four are a good choice [11].

V. CONCLUSIONS AND FUTURE WORK

Given our society’s increasing need for high quality teaching and training, computer-supported education is becoming critical to complementing human tutoring in a large variety of fields and settings. Research in Intelligent Tutoring Systems leverages advances in Artificial Intelligence, Cognitive Science and Education to increase the ability of computer supported education to autonomously provide learners with effective educational experiences tailored to their specific needs, as good human tutors do [4]. Now a days the recent trend is to design e-learning systems that can be made available in various languages (multilingual) it has become much easier to convert the applications/programmes in different languages. India is already multicultural/multilingual country; if the programme can be available in other languages it will benefit lots of learners from various parts of India/World.

It is well advocated that there are two primary advantages of using Semantic web-based model for designing ITS: One is that it contains a hierarchical contents structure and semantic relationships between concepts, can provide related useful information for searching and sequencing learning resources in web-based e-learning systems. The other is that it can help a developer or an instructor to develop a learning sequence plan by helping the instructor understand the why and how of the learning process. A list of the technologies used in the implementation of semantic web-based e-learning system include PHP Platform, Apache Web Server, MySQL database, and RAP Semantic Web Toolkit. As a future work these may also be explored in order to come up with a robust
model of adaptive ITS based on semantic web. Some computer scientists are now striving for simple-minded intelligence that would be easier to develop and perhaps only marginally inferior to full-fledged intelligent tutors. Further advances in information technologies may open other new possibilities.

REFERENCE


