Interactive Ontology-Based User Modeling for Personalized Learning Content Management

Ronald Denaux[1], Vania Dimitrova[2], Lora Aroyo[1]

1 Computer Science Department, Eindhoven Univ. of Technology, The Netherlands
2 School of Computing, University of Leeds, UK
r.o.denaux@student.tue.nl, l.m.aroyo@tue.nl, vania@comp.leeds.ac.uk

Abstract. This position paper discusses the need for using interactive ontology-based user modeling to empower on the fly adaptation in learning information systems. We outline several open issues related to adaptive learning content delivery and present an approach to deal with these issues based on the integration of two existing systems - AIMS (task-based information retrieval environment) and STyLE-OLM (interactive open learner modeling tool). The work contributes to achieving semantic-based reasoning for educational systems and shows a promising direction for the implementation of personalized educational semantic web.

1 Introduction

In the past decade we have witnessed a growing interest in applying personalization and adaptation in numerous application domains. Advanced information systems are styled in a way that enables users to quickly access information relevant for their needs [1]. Deployed in educational settings, these systems will enable offering the most appropriate resources tailored to the learners' needs.

There are two key questions in this context. Firstly, attention should be paid to the use of *technology for engineering of large amounts of domain information and resources* to appropriately tailor them to the individual preferences and knowledge state of different users. The current effort in Semantic Web research examines means for employing ontologies to achieve semantically rich, well-structured, standardized and verified content [2]. The engineering of adaptation and personalization will have to take into account the ontologies used to represent learning content. It is important to understand that adding some semantics to resources has a push effect. This enables the system, and its creators, to have control over what resources are given to the learner and, hence, to shape his/her learning. Thus, appropriate methods are required to tailor the "push" to the needs of each individual user.

Secondly, in order to retrieve the most appropriate learning objects *user semantics should be captured and used.* Typically, adaptive systems maintain a *user model* of the individual user as an overlay of the domain model in order to record the current state of the user with respect to his/her knowledge of domain concepts [3]. There are several aspects to consider: (1) initialization and gradual maintenance of the user model; (2) quick and most relevant access point to the user; (3) accuracy of the system's assumptions about the user's knowledge, goals, task, etc.; (4) empowerment of the system's diagnosis with learner input; and (5) dealing with information not encoded in the system ontology.

In this position paper we argue that the above issues can be tackled by appropriately combining recent approaches from student modeling and adaptive information retrieval. We discuss advantages of integrating interactive ontology-based student modeling and adaptive learner content management to enable offering information resources tailored to the needs of individual learners. Two systems, developed by us in previous studies, are used to show how their integration enables rapid learner modeling that captures the dynamics of the learner's knowledge and ensures effective personalized learning content management.

Next in the paper we will present the need for integrating ontology-based learner modeling into adaptive learning content management systems and will point at issues that have to be tackled (Section 2). We then show in Section 3 how adaptive learning content management (illustrated in AIMS) [4] and interactive ontology-based student modeling (illustrated in STyLE-OLM) [5] may be used to address the open issues. We argue that the integration of these independently developed systems, as presented in Section 4, will enable capturing the user's view of domain ontology and course structure which will empower the personalized learner support provided by AIMS. Finally, our work will be put in a larger context and future plans will be sketched out.

2 The Context

In this section we will outline the context of our research with the help of three user scenarios. We consider a situation where a learner has to explore materials from a large repository in order to accomplish a learning goal. Each learning object is linked to a list of concepts that, according to the beliefs of the creator of the object, will be mastered when the learner studies the object content. For the sake of clarity, we will use here concepts from the topic *Learning about Linux* which includes main concepts needed to understand how the Linux operating system functions (domain concepts are shown in `typewriter`).

2.1 Scenario 1: The Cold Start Problem

Bill is a Computing student and has an assignment that requires using the `Linux operating system`. Bill happens to have previous experiences with the `Microsoft Windows` and `MS-DOS operating systems` and already knows about `operating systems`, `file systems`, and `command-line interfaces`. Bill has been advised to use a web-based system that provides learning materials on Linux. His goal is to log to the system and quickly learn how to copy his files from a floppy disk and how to compress them, as required in his assignment.

Some adaptive systems would allow Bill to directly search for concepts of interest or to choose a topic to study from a list of contents [6]. In this case, the system cannot accurately diagnose Bill's goals and prior knowledge based on his choice. The system may provide learning objects confusing him as they may contain unfamiliar concepts. Furthermore, Bill's goal of working with learning objects may not necessarily correspond to the goal defined by their creators.

In other adaptive tutoring systems Bill may have to do a predefined test allowing the system to extract a model of his domain knowledge [6]. This may have a negative effect on Bill's learning motivation because of both the unpopularity of the tests and the learning distraction that testing may cause.

Now imagine a content management system which first *interacts* with Bill in the form of a short dialog to determine his previous knowledge and how it relates to his learning goal. To be domain independent, the interaction should be guided by a generic dialog planning mechanism linked to reasoning about the course domain and structure. As a result the system will extract an *initial learner model* which will determine an adequate starting point for Bill and will further recommend appropriate learning objects.

2.2 Scenario 2: Semantics is Defined Solely by a Course Creator

Linus, as opposed to Bill, has quite some experience with Linux, but he also wants to use resources to learn more about Linux. Suppose now that both Bill and Linus want to learn more about `user accounts` and `file security`.

Some adaptive systems would offer different courses, e.g. for beginners, intermediate, and advanced learners. Stereotyping in learning systems is prone to problems. It is difficult to define a general meaning for each of the above categories, let alone to correctly classify a learner into one. However, the main drawback is that it is almost impossible for the course creator to consider the specific goals of each learner and to determine the concepts he will refer to when trying to understand a topic within each category.

What one ideally wants is the system to use a user model in *combination* with the domain, course and resource knowledge to attend the needs of each learner as best as possible. Choosing resources will *depend on the learner's semantics* of what the goal is and what concepts are relevant rather than on the semantics pre-encoded by experts unaware of the specific needs of each individual learner.

2.3 Scenario 3: Inaccuracy and Dynamics of the Learner Model

Assume that both Bill and Linus have spent considerable time with an adaptive course management system. Based on some user modeling mechanisms, e.g. tests or tracking the user's interactions with learning objects, the system builds a model that represents the user's understanding of the domain [6].

The diagnostic mechanism may produce inaccurate learner models. For instance, the correct answers in tests may be guessed and may not necessarily be based on domain understanding. In addition, the system's judgment about the learner's understanding based on pre-encoded ontology and on reading a document, may often be inaccurate. For example, let us suppose that Bill and Linus both have read a document about `user accounts` that includes other domain concepts, such as `file system`, `root`, `file permissions`, etc. Bill may be able to create a link between `user account` and `file permissions` but may be unsure of all relations between these concepts, while Linus, who already has some knowledge of both concepts, will be able to understand deeply what their relations are. Reading the document, Bill may also discover the link between `file system` and `root`, although that link is not explicitly discussed in the document on `user accounts`. Linus, being confident of his knowledge on `user accounts`, may skim through the document and, thus, miss some elaborated domain concepts' relations but the system may not have any means to account this.

Furthermore, the learner modeling mechanism may not be able to capture the dynamics of the user's knowledge. For example, Bill and Linus may use other resources to learn about Linux. They may test what they have learned in practice, and may compare their experience with that of fellow students. Their learner models will evolve: more beliefs will be acquired and existing ones will be refined. It will be impossible to capture the dynamics of the learner's knowledge with conventional diagnostic mechanisms.

A desired situation would be to have some combination of *open learner modeling* [7] and *guided diagnostic interaction* [5]. For example, after a knowledgeable learner (like Linus) finishes with a learning object, a clarification dialog can be initiated to enable the learner to identify learned concepts and discovered domain relations. For less knowledgeable learners (like Bill) the clarification dialog may be triggered by showing them their learner models, as discussed in [5].

2.4 Open Issues

The scenarios presented above highlight the following open issues:

- **Dealing with cold start**. Adaptive systems face problems to deal with first time users for whom a user model is not accumulated yet. Rapid user modeling techniques are needed.
- **Use of previous knowledge and experience**. The same document can have different learning effects on different people depending on their prior knowledge. Methods for diagnosing the user's prior knowledge are needed.
- **User models based solely on analysis of learner-system interaction are inaccurate**. In real life, teachers use implicit ways to assess the students' results and performance. Learning systems need explicit mechanisms to clarify the obtained learner models.
- **Semantics of the course author and the learner may differ**. This leads to inaccurate learner diagnosis or inappropriate recommendation of relevant material. Furthermore, such difference justifies why one cannot rely only on the learners' opinion of their domain knowledge (as in open learner models [7]) and that some clarification dialog is needed.
- **The course author's goals and the learner's goals may differ**. Most educational systems assume these two goals are equal and do not adapt to special goals the learners might have. Mechanisms for clarifying the users' goals are critical for providing effective recommendations and guidance.

– **The learner's goals, preferences, and knowledge evolve**. The changes in the learner model are not only gradually between the start and end of the course, but also temporarily during a certain session. Appropriate mechanisms for capturing this dynamics should be employed.

We argue that a fruitful approach to address the above issues is the combination of interactive user modeling and adaptive educational content management.

3 Possible Technologies

This section will briefly outline two existing systems focusing on how they tackle some of the open issues, and will justify the need for their integration.

3.1 AIMS: Adaptive Learning Content Management

AIMS is an *information handling support system* focused on an efficient information provision for *task-oriented problem solving*. It offers adaptive contextual support (through its Search and Browse Tool) that enables users to identify information necessary for performing a particular learning task [4]. As a consequence of the task-orientation, AIMS focuses on the classification of the knowledge in the application (application semantics). Thus, concept-based representations of the content play a pivotal role in modeling the domain content (*domain model*) and the user (*user model*), but also in modeling the application problem-solving strategies (*task model*) and the adaptive mechanisms employed in the system (*sequencing model*). The strict separation between domain-dependent and application related issues and the resources themselves, allows for flexible solutions and reusability. It offers a good content/information basis for adaptive recommendations (as illustrated in the three scenarios), where the domain and the sequencing strategies can be changed easily in order to achieve adaptation.

Adaptive task-based systems like AIMS appear to be especially effective in learning and training oriented application areas, where the sequencing of content, and the concept-based visualization of search results and domain concepts, proves to be a good instrument to guide the user through the material. A key role for achieving this adaptation is played by the user model (its construction and use by both the system and the learner).

AIMS is an example of an adaptive system that suffers from open issues outlined in Section 2.4, e.g. cold start problem, lack of alignment between author's semantics and learner's semantics and lack of understanding of the user's goals (and adapting to them). There is a need for empowering AIMS with user modeling approaches to allow for rapid user modeling, accuracy in the user assessment and alignment of the user's goals and semantics with the ones of the course author. Some of the aforementioned problems are addressed in STyLE-OLM.

3.2 STyLE-OLM: Interactive Ontology-Based Student Modeling

STyLE-OLM is a learner modeling system where a learner model is constructed with the active participation of the learner being allowed to inspect and discuss the content of the model the computer builds of him/her [5]. The system is based on a computational framework for interactive open learner modeling that includes a *domain ontology*, *discourse model*, and belief mechanism for *maintaining a jointly constructed user model*. The framework is fairly general and fully domain independent, two instantiations of STyLE-OLM in a Computing and Finance domain have been demonstrated. STyLE-OLM deals with an *extended overlay learner model* that captures not only correct and incomplete learner beliefs (i.e. overlay upon the domain ontology) but also erroneous beliefs (that are not confirmed by the domain ontology). Patterns of the learner's reasoning that might cause erroneous beliefs are also modeled.

The interaction in STyLE-OLM is in a graphical manner and has two modes: DISCUSS, where the learners discuss aspects of their domain knowledge and influence the content of their models, and BROWSE, where the learners inspect the current state of their models. Throughout the discussions, the system makes plausible inferences about what is further believed by the learner on the basis of what is explicitly asserted, and from this a *dialog strategy* is determined. Important in the context of this paper is that STyLE-OLM can be adapted by changing its dialog strategies to tackle knowledge alignment issues, e.g. by following studies about differences that may occur between conceptual models of people [8]. Knowledge probing strategies could also be added to articulate the learner's conceptualization at the beginning of a session.

STyLE-OLM and AIMS are complementary, their integration, discussed below, allows us to address issues outlined in Section 2.4.

4 Integration of AIMS and STyLE-OLM

There are several opportunities for integrating features of both systems (see Figure 1). As pointed in Section 3.1, the architecture of AIMS has to be extended with an appropriate user modeling component. STyLE-OLM is the kernel of this component, as it provides a dialog-based interaction with the user and is a rapid way for extracting information about the user.

4.1 The Role of Ontology in the Integrated Architecture

Both STyLE and AIMS have used ontologies developed for their specific needs. The use of ontologies, especially if they are based on standards, provide several advantages: *extensibility and flexibility* (we can plug in other ontologies or relate our domain ontology to a top ontology), *interopability* (we can share ontologies with other applications) and *reusability of previous work* (we can use existing tools and methods for eliciting, aligning, parsing and visualizing ontologies). A stimulating discussion on these issues in a broader scope is given in [9]. AIMS has already been moved towards supporting standard ontology languages like OWL. STyLE-OLM is being adapted for dealing with a domain ontology encoded in OWL, as well as for maintaining a user model represented in OWL. Using a *common ontology encoded in OWL* is critical and will ensure interoperability of

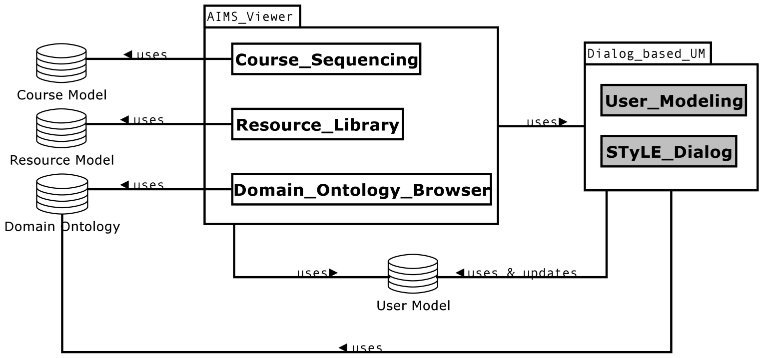

Fig. 1. Integrated architecture of AIMS, STyLE-OLM and user modeling component.

components of both systems and wide application of the integrated system. The other key aspect of the integration is the *common user model*, discussed below.

4.2 Information Related to the User Model in AIMS

Each of the information models (domain, resource library and tasks) of AIMS has some relation to the user model. The domain model provides a point of reference to be used for building an extended overlay learner model. The resource library model is related to the user model by keeping track of which resources have been used by the learner and measuring the learning effects of the resources on the learners. The tasks model in AIMS can be a source of meta-level information about the user and the domain, as it provides the following information:

- **Purpose of the current task**. Each task in AIMS has a purpose which corresponds to the goal the teacher has in mind for the student when performing the task. In practice, the goals of the student might differ from the intended purpose of the course author. Examples of purposes are: Learning new concepts, Brushing current knowledge about some concepts, Reviewing some topics which are already known, etc.
- **Focus of the task**. Each task in AIMS is related to a set of concepts X which the learner should learn about. It can be used to maintain the learner's focus on the current task.
- **Ideal or standard learning path**. The tasks model contains information about pre-conditions and learning effects of tasks. This information can be used to infer which concepts are basic knowledge for the course (concepts required by most of the tasks) and which are intermediate or advanced knowledge. This information could be used to try to solve the cold start problem by using STyLE-OLM to determine the learner's knowledge about concepts required for tasks.

4.3 Analysis of the User Model in AIMS with STyLE-OLM

The previous paragraphs discussed how the current information in AIMS about domain, course and resources can be related to user modeling information. We will now discuss how the user model itself can be used in AIMS and how STyLE-OLM can be used to extract parts of the user model and as an interface between the system and the user.

- **Validating task and resource effects**. Right now, AIMS assumes that completing a task or viewing an educational resource will result in the same learning effect for every user. STyLE-OLM can be used as a review of the task or resource to find out what the learner's knowledge is after completing the task or viewing the resource. Depending on the desired quality of the UM, this STyLE-OLM interaction could be done after every task or after a series of tasks or only when a user requests such a review.
- **Tackling the cold start problem and aiding with the task sequencing**. STyLE-OLM could use the tasks model to determine which concepts to probe for if no information is known about a user. In general, STyLE-OLM could be used to determine whether a user has enough knowledge for recommending a task.
- **Adaptive browsing**. AIMS can narrow down the number of concepts shown. The current task already roughly defines which concepts should be learned by the user during the task. By combining the user knowledge, preferences and goals and the specification of the current task, it is possible to determine which concepts and links are more relevant when browsing.
- **Visual representation of open learner model**. The idea behind this is to show the user what their model is. This has two advantages: (1) the user can disagree with their model and initiate a dialog with STyLE-OLM to discuss about the concepts (as a result, the learner model will be refined); (2) the user can see which concepts are little known to him/her and can decide to learn about them.
- **STyLE-OLM as an aid for understanding the domain**. When browsing the domain, users can use STyLE-OLM to focus on a certain set of concepts and discuss their understanding about those concepts. The students can have the initiative every time when starting a discussion and STyLE-OLM will answer students' questions or help the students find out what are the wrong assumptions/conclusions they make.
- **Using the UM for selecting and sorting resources**. When users want to learn more about a certain topic, they can query the resource library to find suitable educational resources. When there are several resources as a query result, these resources have to be filtered and sorted. This can be done based on the user's current knowledge, goals and preferences to determine the kind of resource most suitable for the user. Examples of preferences are preferences for a kind of material format (multimedia or only text) and preference for practical oriented resources (tutorials and examples instead of theoretical resources).

– **Articulating the learner's goals**. To be able to adapt to the user's goals we need to identify those goals. For this we may use information from the current task, a STyLE-OLM dialog or a form filled in by the user. In order to provide this kind of adaptation we need enough and appropriate annotation of the resource library. Also, because the user's goals, preferences and knowledge change in time we need a way of modeling users through time in order to maintain continuity and momentum in the learning activities. We consider the use of *episodic* and *long term* user models. During dialog episodes, a short-term, episodic user model is accummulated. It is used as a source for updating the long term user model at the end of each episode by employing STyLE-OLM algorithms based on non-monotonic reasoning.

5 Conclusions

In this paper we have analyzed the problems of current adaptive educational information systems, and have stressed the need for deploying advanced user modeling techniques to achieve more adequate personalization and adaptation to the individual user needs. We have proposed an integrated approach based on ontologies to represent both domain and user knowledge, and to maintain a user-system dialog that initializes and manages the user model. The work contributes to achieving ontology-based reasoning for personalizing the educational semantic web. There are several ongoing projects in this increasingly important research field. For example, in a recently started project McCalla, Greer, Vassileva and colleagues are employing methods from their work on distributed student modeling to attach user models to learning objects in order to obtain information for personalizing the use of educational resources [10]. Judy Kay and colleagues are adapting scrutable user modeling approaches to capture a student's understanding of ontology [11]. Simon, Dolog and colleagues discuss the use of semantic web technologies in a broader context for learner modeling, learning resource modeling and its matching for providing adaptation [9]. Our work offers an alternative approach that builds upon previous methods for interactive open student modeling and adaptive learning content management. We believe that this is a promising direction that will contribute to the implementation of personalized educational semantic web.

As a next step we aim at designing a computational framework for the use of semantic web technologies (focusing on domain ontology but also considering other types of ontologies) to enhance the effectiveness of adaptive learning systems (considering adaptation both to learners and course authors). We intend to illustrate the framework with the integration of AIMS and STyLE and perform series of user studies in order to assess the effectiveness of the approach. This work is part of the SWALE (Semantic Web and Adaptive Learning Environments) project, started in April 2004, and funded by the British Council and NWO. SWALE is targeted to: (1) identify dimensions of knowledge sharing, employed currently in semantic web technologies, that can be applied to learning systems to enable different viewpoints (contexts) on learning material; (2)

determine conditions which facilitate the construction of a shared ontology in learning systems to support the joint construction of educational materials and adaptation to individual teacher/learner needs; (3) develop methods to formalize the construction of shared knowledge in a pedagogical context to provide effective author guidance and learner feedback in environments for individualized learning and course delivery.

Acknowledgments. This work contributes to and is supported by the PROLEARN EU network of excellence, specifically the work package on Personalized Adaptive Learning. The research is partly supported by the UK-Netherlands Partnership in Science programme. Special thanks to Prof. Paul de Bra for helping with the preparation of the SWALE project and supporting Ronald Denaux's visit at Leeds. We thank members of the Adaptive Hypermedia group at Eindhoven and the Interactive Learning Environments research group at Leeds for many helpful comments and stimulating discussions.

References

1. Brusilovsky, P., Tasso, C.: Special issue on user modelling for web information retrieval. User Modeling and User Adapted Interaction **14** (2004)
2. Ding, Y., Fensel, D., Klein, M.C.A., Omelayenko, B.: The semantic web: yet another hip? Data Knowledge Engineering **41** (2002) 205–227
3. Brusilovsky, P.: Methods and techniques of adaptive hypermedia. User Modeling and User-Adapted Interaction **6** (1996) 87–129
4. Aroyo, L., Dicheva, D.: Aims: Learning and teaching support for www-based education. Int. Journal for Continuing Engineering Education and Life-long Learning (IJCEELL) **11** (2001) 152–164
5. Dimitrova, V.: Style-olm: Interactive open learner modelling. Int. Journal of Artificial Intelligence in Education **13** (2003) 35–78
6. Brusilovsky, P., Paylo, C.: Adaptive and intelligent web-based educational systems. Int.Journal of Artificial Intelligence in Education **13** (2003) 159–172
7. Morales, R., Pain, H., Bull, S., Kay, J., eds.: Workshop on Open, Interactive,and Other Overt Approaches to Learner Modelling, at AIED99 (1999)
8. A. Hameed, D. Sleeman, A.P.: Detecting mistmaches among experts' intologies acquired through knowledge elicitation. Knowledge-Based Systems **15** (2002) 266–273
9. Simon, B., Dolog, P., Miklós, Z., Olmedilla, D., Sintek, M.: Conceptualising smart spaces for learning. Journal of Interactive Media in Education (2004(9))
10. McCalla, G.: The ecological approach to the design of e-learning environments: Purpose-based capture of use of information about users. Journal of Interactive Media in Education (2004(7))
11. Cimolino, L., Kay, J.: Verified concept mapping for eliciting conceptual understanding. In Aroyo, L., Dicheva, D., eds.: Proceedings of ICCE Workshop on Concepts and Ontologies in Web-Based Educational Systems. (2002) 11–16