A Framework for a Multi-Faceted, Educational, Knowledge-Based Recommender System

John W. COFFEY
Department of Computer Science
The University of West Florida
Pensacola FL 32514, USA

ABSTRACT

The literature on intelligent or adaptive tutoring systems generally has a focus on how to determine what resources to present to students as they make their way through a course of study. The idea of multi-faceted student modeling is that a variety of measures, both academic and non-academic, might be represented in student models in service of a broader educational context. This paper contains a framework for a multi-faceted, educational, knowledge-based recommender system, including a basic set of descriptors that the model contains, and a taxonomy of inferences that might be made over such models.

Keywords: recommender system, intelligent tutoring system, student model, Semantic Web

1. INTRODUCTION

Student modeling is an endeavor within the broader realm of user modeling that seeks to create representations of various attributes of students that can be utilized through computerized means to enhance the educational process. Most modern work on such systems focuses on adaptive lessons that respond to student progress by altering lesson content. Work in the broader area of student modeling to recommend learning resources tailored to individuals or collaborating groups, or through recommendations regarding the constitution of the groups themselves also holds promise. In addition to the representation of the results of academic attainment testing, such a model might record the results of personal inventories that identify instructional preferences, social interaction modes, learning style, etc., as part of the modeling process.

The resulting multi-faceted student model could be used to assist in making inferences regarding resources that might be of interest to students, to recommend alternative ways to organize groups, or to make remediation decisions or recommendations. The impetus for this paper comes from an effort to create a framework for multifaceted student modeling systems in service of a variety of goals, including the recommendation of learning resources to individual students, inference over a range of student attributes to construct collaborating groups, recommendations of resources to collaborating groups, and organization of multiple resources for individuals or groups, based upon attributes of the resources and groups.

The rest of this extended abstract contains a brief review of pertinent literature regarding student modeling and personalization to facilitate learning, and a framework for multi-faceted student modeling that addresses the types of inferences that might be made over such models. The framework includes people-to-people relationships, resource-to-resource relationships, and people-to-resource relationships.

2. RELATED LITERATURE

Several basic questions underlie this work. For instance, what is the range of attributes that should be included in a student model? Given that one has a student model, what might be done with it? The next two sections address the structure of student models, the types of attributes that might be included in a student model, and various purposes for which student models might be employed.

Student Modeling

Researchers have explored and refined methods to create and maintain computer-based user models for years [1, 2], and the literature contains descriptions of both implicitly and explicitly constituted models. Implicit user models are those that are created on-the-fly from user interactions with the system that is creating the model. Explicit user models are created through the elicitation of information directly from the user. This section contains descriptions of representative literature on both implicit and explicit student modeling that spans some number of years, in service of elucidating the range of features that might be modeled.

Barker, Jones, Britton and Messer [3] describe a comprehensive set of student attributes that fall into four broad categories: language level, cognitive style, task and question levels, and help level. Student models are created both by assigning values automatically (implicitly) and through "cooperative configuration" (explicitly) of variables representing the various elements of the model. The models are used to decide which version of multimedia instructional materials are presented to the students, a form of student stereotype approach [1] that requires the ability to place students into stereotypic categories. The addition of goals and interests to the items modeled by Barker et al. would create a more comprehensive set of student model attributes.

Bianchi-Berthouze and Lisetti [4] described an explicit approach to the acquisition and modeling of affective, subjective experiences. They describe a Model of User Emotions (MOUE) that employs sensors for facial expressions, and Multimedia Interactive Environment for Kansei (MIKE), which uses verbalizations to form its model of the user's affective state. If well realized, such sensing systems might prove useful in learning systems, by enabling the system to assess whether or not the user is satisfied with the course of events that are unfolding. However, since these systems rely on feedback loops to ascertain the accuracy of their assessments of affect, the possibility of the system being overly intrusive is an issue.

Li and Yoo [5] describe an adaptive online tutoring system named AToL that assesses student learning style and then employs a Markov chain to cluster interactions that become a model of student behavior. Piech et al [6] describe the use of machine learning to model student performance as they solve introductory programming problems. They claim that their method can be used to predict which students will do well and which will have difficulty mastering introductory programming. Ostrow et al [7] describe the use of student models for adaptive learning systems. They describe a way to assign partial credit to weigh problem difficulty for the next problem offered to the student in an intelligent tutoring system. Allen, Snow, and MacNamara [8] describe a special purpose student model of reading ability and an intelligent tutoring system named iSTART to teach reading. iSTART uses natural language processing to build comprehension models from student explanations of the content they have read.

Personalization to Facilitate Learning

A substantial body of literature on personalized learning systems forms the basis of the literature regarding uses of student models to support learning. Brusilovsky [11] describes various personalization elements such as adaptive navigation, recommendations regarding where to go next, and adaptive presentation, in which the system can build a sequence of items for the learner to peruse, or recommend items that might be of interest. Burke [9] enumerates several categories of recommender systems including content-based, collaborative filtering, knowledge-based, and demographic. She states that knowledge-based recommenders offer a variety of benefits including the ease with which new users can use the system, sensitivity to changes in student preferences, the ability to include knowledge relevant to a variety of features, and the ability to create a mapping from student needs to beneficial learning materials.

Calvi and Cristea [10] state that adaptive systems for education typically employ basic rules that only address presence or absence of prerequisites, and propose an extended taxonomy of rules such as level rules, temporal rules, repetition rules, and rules to deal with generalizations and specializations of content. As examples, level rules essentially seek to capture whether or not the learner has enough prerequisites or has spent enough time spent on precursor topics before advancing to the next topic. Generalization and specialization rules allow for inductive reasoning from a current item of interest to more general items, and deduction to more special-case ones.

3. A FRAMEWORK FOR AN EDUCATIONAL RECOMMENDER SYSTEM

The following section contains a description of the framework for an educational, knowledge-based recommender system. As stated, knowledge-based recommenders can be rule-based, case-based, etc. The system described in this article is a system that utilizes attainment data (as described next), along with a variety of other relevant information such as learning style and demographic data to make a range of recommendations.

Descriptions of the types of student attributes that are modeled in this system, and a taxonomy of inference rule types that might be formulated as part of the recommendation process are presented here. A number of queries that illustrate various aspects of the model have been prototyped in both SWI Prolog and in the SPARQL query language [14] for the Semantic Web. Examples of typical SPARQL queries are presented in Section 5.

Attributes of the Student Model

Figure 1 contains a basic representation of student model attributes and rule types that are employed in this framework. Student attributes pertaining to academic performance and non-course measures are part of the student modeling system.

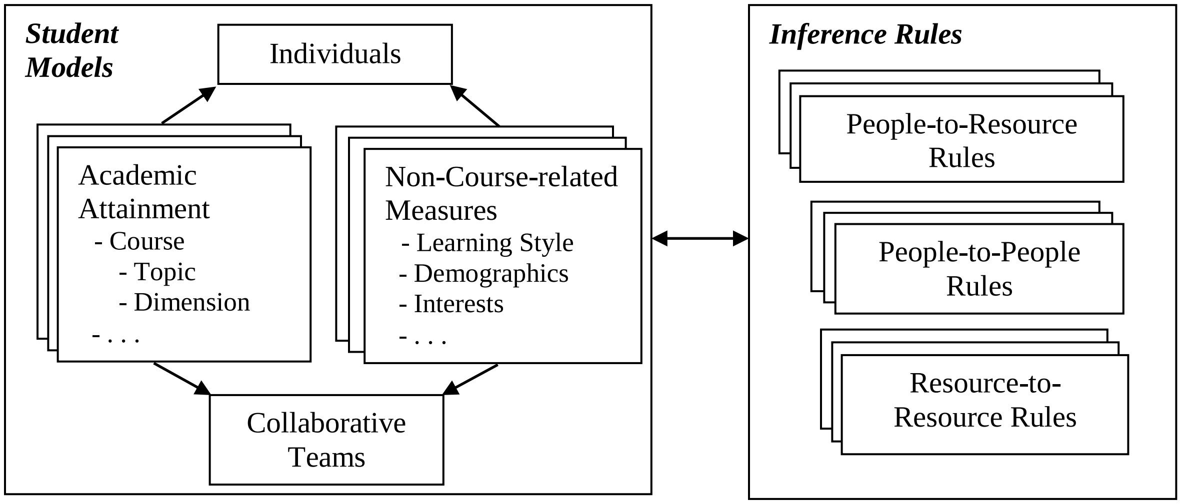

Figure 1. A representation of the modeling framework.

A hierarchy of elements pertaining to courses and topics within a course are captured in the academic attainment component. A number of topics are aggregated together as courses. Assessments of performance in individual topics are based upon a number of dimensions that may be particular to courses or the individual topics. Other types of academic attainment measures capture attributes including reading level and mathematical competency.

A number of non-course-related attributes are included in this model. Examples include elements such as learning style, interests, past individual/group projects, professional or experiential learning, etc. Additionally, demographic information such as age, gender, and location are of use for the creation of heterogeneous learning groups. Learning style information is used to develop teams of students with similar or complementary approaches to problem solving and learning. For instance, a continuum of the Felder Inventory of Learning Styles [13] is the active-reflective continuum. Active learners like to start immediately doing things whereas reflective students think before acting. Combining students with complementary learning styles is a goal.

A Taxonomy of Inference Categories

Inference over student models are in service of a variety of goals. The current framework is comprised of three non-disjoint general categories of inference, which are:

- People-to-people
- Resource-to-resource
- People-to-resource

People-to-people inferences are used by themselves to identify potential collaborators in the formation of work groups. Resource-to-resource queries are used to assemble packages of resources to address specific learning objectives. People-to-resource queries are used to match individuals or collaborating teams to resource packages.

Inferences are made regarding the formation of collaborating teams of students to work groups with similar or complementary approaches to studying and learning. Other sorts of people-to-people inferences involve identifying those with shared interests, shared deficiencies to remediate, shared or complementary demographic values (placing some older students with some younger students, creating more diverse groups, finding collaborators who live in the same area), etc. Resource-to-resource queries utilize attributes of resources in order to build aggregations of resources for individuals or groups. Basic attributes of resources include content area (to what topic does the resource pertain?), the basic type of document: theoretical, applied, case study, etc., intended audience: introductory, intermediate or advanced, reading level, and chronology of document formation (to address the evolution of thought in the area).

Descriptors are used to create sequences of resources. For instance the system might suggest that a theoretical description be viewed first, followed by a simple case study that makes sense within the context of the theory, followed by a more complex case study that requires generalization of the theory. Tailoring appropriate resources to the intended audience

```
<?xml version='1.0'?>
<rdf:RDF
    xmlns:leos='http://www.coginst.leo.edu/LEOStudentsOnt#'
    xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
    <rdf:Description rdf:about="leos:Joe Smith">
        <leos:StuNum>12345</leos:StuNum>
        <leos:ZipCode>32561</leos:ZipCode>
        <leos:Score>60</leos:Score>
        <leos:active-reflective>Act</leos:active-reflective>
        <leos:global-sequential>Glo</leos:global-sequential>
        <leos:interested_in>Poetry</leos:interested_in>
        <leos:Learninggoal>Global</leos:Learninggoal>
   </rdf:Description>
   <rdf:Description rdf:about="leos:Seeking revenge">
        <leos:atheme>Tragedies</leos:atheme>
        <leos:atype>Seeking revenge</leos:atype>
   </rdf:Description>
   <rdf:Description rdf:about="leos:Ambition">
        <leos:atheme>Tragedies</leos:atheme>
        <leos:atype>Ambitious</leos:atype>
   </rdf:Description>
   <rdf:Description rdf:about="leos:Ambition">
        <leos:described_in>http://www.coginst.leo.edu/Macbeth's Ambition
        </leos:described_in>
   </rdf:Description>
   <rdf:Description rdf:about="leos:Seeking revenge">
        <leos:described_in>http://www.coginst.leo.edu/Hamlet/Hamlet's Flaw
        </leos:described_in>
   </rdf:Description>
 </rdf:RDF>
```

Figure 2. Knowledge Representation utilizing standard and domain-specific namespaces.

(introductory, intermediate or advanced) is important. Finally, consideration of reading level and chronology of document formation may all be utilized by the system. A prototype system that implements the framework has been created. Capabilities from the Semantic Web including the use of XML, RDF and the SPARQL query language [14] have been used. Typical components of the prototype are described in the next section.

4. KNOWLEDGE REPRESENTATION

Figure 2 contains a small excerpt from a knowledge model of the system. Basic information about a student is represented including name, student number, zip code, scores on learning activities and learning style. Additionally, metadata about the student's interests and typical semantically marked-up learning resources is included.

The knowledge representation utilizes the standard Resource Description Framework (RDF) namespace. Namespaces are used to disambiguate names. For instance, if one saw mention of a table, it might not be clear if the reference were to a physical entity usually having four legs or to an Excel worksheet.

The knowledge representation also contains a custom namespace (`leos`) that has been created to model people and resources in support of the framework described in this work. The concepts in this small sample are completely generalizable to any student or resource description.

5. EXAMPLE QUERIES

Figure 3 contains two rules that provide examples of a people-to-people query and a people-to-resource query. Two classes of such rules have been created:

those that pertain to recommendations for individuals and those that pertain to recommendations for groups. Query 1 in Figure 2 presents an example of a rule that pairs two students who share similar interests, and the query recommends a work that addresses that interest. It is, however, fundamentally a people-to-people query.

In Query 2, the system matches students who are interested in a particular category of project. It then identifies various descriptors of projects in that general category and finds specific projects that match those descriptors. Finally, the query seeks to match students with complementary learning styles and who live close together. Arbitrarily simple or complex combinations of student academic and non-academic attributes may be combined by extending the prototype system.

While more empirical research is needed on comprising teams of students with complementary learning styles, a strong intuition exists that good teams might be made of combinations of active and reflective students with the active students encouraging the reflective ones to experiment and the reflective students encouraging the active students to think more before they start to tinker.

While development of the Semantic Web has progressed more slowly than originally thought, semantically marked-up student attributes and learning resources hold great promise to facilitate a variety of educational goals in the future. The framework outlined here addresses many of the attributes and issues such a system might embody. Future work will include use and testing of the prototype.

```
SELECT ?studentx ?studenty ?yInterest ?aType ?aWork
WHERE (?studentx, <leo:LEOOnt#interested_in>, ?xInterest),
      (?studenty, <leo:LEOOnt#interested_in>, ?yInterest),        (1)
      (?aType, <leo:LEOOnt#atheme>, ?xInterest),
          (?aWork, <leo:LEOOnt#described_in>, ?attribute)
AND ! (?studentx eq ?studenty)
AND (?xInterest eq ?yInterest)
AND (?aType eq ?attribute)

SELECT ?studentx ?studenty ?xAR ?xGS ?yAR ?yGS
WHERE (?studentx, <leo:LEOStudentsOnt#active-reflective>, ?xAR),
      (?studentx, <leo:LEOStudentsOnt#global-sequential>, ?xGS),  (2)
      (?studentx, <leo:LEOStudentsOnt#ZipCode>, ?xZip),
      (?studentx, <leo:LEOStudentsOnt#Score>, ?xScore),
      (?studenty, <leo:LEOStudentsOnt#active-reflective>, ?yAR),
      (?studenty, <leo:LEOStudentsOnt#global-sequential>, ?yGS),
      (?studenty, <leo:LEOStudentsOnt#ZipCode>, ?yZip),
      (?studenty, <leo:LEOStudentsOnt#Score>, ?yScore)
AND ! (?xAR eq ?yAR && ?xGS eq ?yGS)
AND (?xZip eq ?yZip)
AND (?xScore < 80)
AND (?yScore < 80)
```

Figure 3. People-to-people and people-to-resource SPARQL rules.

4. SUMMARY AND CONCLUSIONS

This article describes a high-level, global framework for student modeling in service of an educational, knowledge-based recommender system. It first presents literature on user modeling that includes a discussion of both implicit and explicit modeling approaches. It then describes a basic student model comprised of both academic attainment and non-course related descriptors of students. It identifies basic categories of inference rules that might be performed over such a student model. The framework accounts for students working individually or in collaborating groups. It also accounts for the identification of resources that may be useful for individuals or groups based upon interest or the need for remediation. The framework identifies categories of inference that include people-to-people, resource-to-resource, and people-to-resource. The article contains a description of a prototype built with Semantic Web technologies that include XML-RDF and the SPARQL query language.

5. REFERENCES

[1] Rich, E. 1979. User Modeling via Stereotypes. **Cognitive Science** 3. pp 329-354.

[2] Finin, T., and Drager, D. 1986. GUMS: A General User Modeling System. **In Proceedings of the 1986 Canadian Society for Computational Studies of Intelligence** (CSCSI-86). pp 24-30. Montreal, Canada.

[3] Barker, T., Jones, S., Britton, C., and Messer, D. 2002. The Use of Co-operative Student Model of Learner Characteristics to Configure a Multimedia Application. **User Modeling and User-Adapted Interaction.** 12. pp 207-241.

[4] Bianchi-Berthouze, N., and Lisetti, C.L. 2002. Modeling Multimodal Expression of User's Affective Subjective Experience. **User Modeling and User-Adapted Interaction.** 12. pp 49-84.

[5] Li, C., and Yoo, J. Modeling Student Online Learning Using Clustering. **Association of Computing Machinery Southeast conference, 2006,** March, 10-12, 2006, Melbourne, Florida, USA. 1-59593-315. pp 186-191.

[6] Piech, C., Mehran Sahami, Daphne Koller, Stephen Cooper, Paulo Blikstein. 2012. Modeling How Students Learn to Program. **Procedings of SIGCSE'12**, February 29–March 3, 2012, Raleigh, North Carolina, USA. ACM 978-1-4503-1098. pp 153-158.

[7] Ostrow, K., Donnelly, C., Adjei, S., Heffernan, N. 2015. Improving Student Modeling Through Partial Credit and Problem Difficulty. **Proceedings of L@S'15, the Second Annual ACM Conference on Learning At Scale**, March 14–18, 2015, Vancouver, BC, Canada. ACM 978-1-4503-3411-2/15/03. pp 11-20.

[8] Allen, L., Erica L. Snow, Danielle S. McNamara. 2015. Are You Reading My Mind? Modeling Students' Reading Comprehension Skills with Natural Language Processing Techniques. 2015. **Proceedings of LAK '15, the 5th International Learning Analytics and Knowledge Conference,** March 16 - 20, 2015, Poughkeepsie, NY, USA. 2015 ACM 978-1-4503-341. pp 246-254

[9] Burke, R. 2002. Hybrid recommender systems: Survey and Experiments. **User Modeling and User-Adapted Interaction.** 12. pp 331-370.

[10] Calvi, L., and Cristea, A. 2002. Towards Generic Adaptive Systems: Analysis of a Case Study. **Lecture Notes in Computer Science,** vol. 2347: De Bra, P.; Brusilovsky, P.; Conejo, R., eds. Springer-Verlag.

[11] Brusilovsky, P. 1996. Methods and Techniques of Adaptive Hypermedia In P. Brusilovsky and J. Vassileva, eds. **User Modeling and User-Adapted Interaction. Special issue on Adaptive Hypertext and Hypermedia**. 6(2-3). pp 87-129.

[12] Mota, P., Melo, F.S., and Coheur, L. 2015. Modeling Students Self-Studies Behaviors. **Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015),** Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey. pp 1521-1528.

[13] Solomon, B., and R.M. Felder, R.M. 2003. Index of Learning Styles Questionnaire. Online [available] **http://www2.ncsu.edu/unity/lockers/users/f/felder/public/ILSdir/ ilsweb.html**

[14] SPARQL. SPARQL Query Language for RDF. Online Available at **http://www.w3.org/TR/rdf-sparql-query/ Accessed 12/30/2015.**