USER MODELING IN ADAPTIVE HYPERMEDIA EDUCATIONAL SYSTEMS

António Constantino Martins*
const@dei.isep.ipp.pt
Luíz Faria*
Carlos Vaz de Carvalho**
Eurico Carrapatoso***

* Knowledge Engineering and Decision Support Research (GECAD)- Institute of Engineering – Polytechnic of Porto
Porto, Portugal
** Institute of Engineering – Polytechnic of Porto
Porto, Portugal
***Faculty Of Engineering – University of Porto

ABSTRACT
This document is a survey in the research area of User Modeling (UM) for the specific field of Adaptive Learning. The aims of this document are: To define what it is a User Model; To present existing and well known User Models; To analyze the existent standards related with UM; To compare existing systems.

In the scientific area of User Modeling (UM), numerous research and developed systems already seem to promise good results, but some experimentation and implementation are still necessary to conclude about the utility of the UM. That is, the experimentation and implementation of these systems are still very scarce to determine the utility of some of the referred applications.

At present, the Student Modeling research goes in the direction to make possible reuse a student model in different systems. The standards are more and more relevant for this effect, allowing systems communicate and to share data, components and structures, at syntax and semantic level, even if most of them still only allow syntax integration.

KEYWORDS: Adaptive Hypermedia System, E-learning, User Model

1 Introduction

Adaptive Hypermedia (AH) is generally referred as a crossroad in the research of Hypermedia and User Modeling (UM) (Brusilovsky, 1996b; Brusilovsky, 2001; De Bra, 2004a; De Bra, 2003a). An AH system (AHS) builds a model of the objectives, preferences and knowledge of each user and uses it, dynamically, through an Interaction Model and a Domain Model, to adapt its contents, navigation and interface to the user needs. The global architecture proposed
by Benyon (1993) and De Bra (2004b), (Figure 1), indicates that AH System (AHS) must have three essential parts:

- The User Model, that describes the information, knowledge, preferences, etc., of the user. This component allows extracting and expressing conclusions on the user characteristics.
- The Domain Model represents a set of domains concepts. In different AHS these concepts can have distinct functions, weights and meanings. Most commonly, each concept is connected / related with other concepts, representing a semantic net. The most important function of this model is to provide a structure for the representation of the user domain knowledge. One approach is to store the estimate level of the user’s knowledge for each concept. This value can be expressed quantitatively, qualitatively or in probabilistic form.
- The Interaction Model, which represents and defines the interaction between the user and the application. The data stored in the Interaction Model are used to infer user characteristics with the objective updating and validate the UM. For that purpose, this component includes evaluation, adaptation and inference mechanisms.

![Figure 1. AHS Architecture (De Bra, 2004b)](image-url)
2 User Modeling

The beginning of User Modeling is dated to 1978/1979 with the first work by Allen, Cohen, Perrault and Rich (Kobsa, 2001). In the following 10 years, numerous applications or systems were developed to store different types of user information to allow distinct adaptation models. Morik, Kobsa, Wahlster and McTear present an extensive survey of these systems (Kobsa, 2001). In these initial systems, user modeling was embedded and there was not a clear distinction from other components of the system (Kobsa, 2001).

In middle 80’s, this separation was made, but no efforts were carried out to allow the reuse of information between adaptive systems (Kobsa, 1985). In 1990, Kobsa was the first author to use the term "User Modeling Shell System". Since then, different systems have been developed with the ability to reuse User Models (Kobsa 2001).

2.1 Student Model (SM)

In generic AHS, the User Model allows changing several aspects of the system, in reply to certain characteristics (given or inferred) of the user (Brusilovsky, 1995, Brusilovsky, 2001). These characteristics represent the knowledge and preferences that the system assumes that the user (individual, group of users or no human user) has.

In Educational AHS, the UM (or Student Model) has increased relevance: when the student reaches the objectives of the course, the system must be able to re-adapt, for example, to his knowledge (Brusilovsky, 2001; Laroussi, 2001).

The relation between Educational AHS and constructivism is now being addressed since it is believed that students do not simply keep the information in a static and isolated form, but look to cognitively build on top of existing blocks of old knowledge (Laroussi, 2001). In other words, the learning process is more efficient when it is built over previously acquired knowledge and it will be more useful if the student is actively implicated in the learning process (Jonassen, 1991).

A Student Model includes information referring to the specific knowledge that the system judges that the user possesses on the domain, known as the Domain Dependent Data (DDD) (Figure 2). The components of the Domain Dependent Data correspond to the Domain Model with a three-level functionality:

- Task level, with the objectives / competences of the domain that the user will have to master. In this case, the objectives or intermediate objectives can be altered according to the evolution of the learning process;
- Logical Level, which describes the user knowledge of the domain and is updated during the student’s learning process;
- Physical Level, that registers and infers the profile of the user knowledge.
The Domain Independent Data (DID) are composed of two elements: the Psychological Model and the Generic Model of the Student Profile, with an explicit representation (Kobsa, 2001). The psychological data are related with the cognitive and affective aspects of the student. Some studies have demonstrated that the difference between the cognitive capacities and personality aspects affects the quality of some models or styles of interaction (Kobsa, 2001; Carrilho, 2004). These data are more permanent which allows the system to know beforehand which are the characteristics that it must adapt to. (Benyon, 1993, Vassileva, 1998). The data related to the user interests, common knowledge and background are kept in the Generic Model of the Student Profile.

The DID include following aspects: (Benyon, 1993; Kobsa, 2001; Carrilho, 2004):

- Initial user knowledge;
- Objective and plans;
- Cognitive capacities;
- Learning styles;
- Preferences;
• Academic profile (technological studies versus economical studies and management, knowledge of literature, artistic capacities, etc.);
• Age and type of student (Kobsa, 1997).
• Cognitive style (affective, impulsive, etc.) personality aspects (introverted, extroverted, etc.) (Laroussi, 2001).

As expressed before, some of these characteristics are relevant for a determined type of UM and not for others (Brusilovsky, 199a; Brusilovsky, 1996b; Brusilovsky, 2001). Therefore, for each AHS, it will be necessary to define which are the characteristics and relevant parameters of the user to be kept.

In table I, we present the more common characteristics used in the UM.
<table>
<thead>
<tr>
<th>Model</th>
<th>Profile</th>
<th>Characteristics</th>
<th>Descriptions / examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Generic Profile</td>
<td>Personal information</td>
<td>Name, email, password, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demographic data</td>
<td>Age, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Academics background</td>
<td>Technological studies versus economics etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Qualifications</td>
<td>Certificates, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge (background knowledge)</td>
<td>A collection of knowledge translated in concepts. Possibility of a qualitative, quantitative or probabilistic indication of concepts and knowledge acquired for the user</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deficiencies: visual or others</td>
<td>Sees well, uses eyeglasses, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Domain of application</td>
<td>Localization of the user etc.</td>
</tr>
<tr>
<td></td>
<td>Psychological profile</td>
<td>Inheritance of the characteristics</td>
<td>Creation of stereotypes that allow to classify the user</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learning style</td>
<td>Definition of the learning style</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive capacities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Domain Dependent Data</td>
<td>Traces of the personality</td>
<td>Psychological profile (introverted, extrovert, active, etc.).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inheritance of characteristics</td>
<td>Creation of stereotypes that allow to classify the user</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objectives</td>
<td>Questionnaires that allow to determine with objectives the user intends to use the system</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Planning / Plan</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complete description of the navigation</td>
<td>Kept register of each page accessed.</td>
</tr>
<tr>
<td></td>
<td>Domain Dependent Data</td>
<td>Knowledge acquired</td>
<td>A collection of knowledge translated in concepts. Possibility of a qualitative, quantitative or probabilistic indication of concepts and knowledge acquired for the user</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Results of assessment</td>
<td>Data of all the tests, exercises, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Context model</td>
<td>Data related with the environment of the user (resolution of the monitor, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aptitude</td>
<td>Definition of aptitude and the capacity to use the system</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interests</td>
<td>Definition of the interests of the individual with the objective to adapt the navigation and contents</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deadline extend</td>
<td>Long, short or normal stated period</td>
</tr>
</tbody>
</table>

Table 1. Common characteristics in User Modelling
2.2 Implementation of the Student Model

Two different types of techniques are used to implement the Student Model: Knowledge and Behavioral based (Kobsa, 2001). The Knowledge-Based adaptation typically results from data collected through questionnaires and studies of the user, with the purpose to produce a set of initial heuristics. The Behavioral adaptation results from the monitorization of the user during his activity.

The use of stereotypes classifies users in groups and generalizes student characteristics to that group (Knobs, 2001; Knobs, 1997, Knobs, 1993). The definition of the necessary characteristics for the classification in stereotypes must take into consideration the granularity degree wanted.

The Behavioral adaptation can be implemented in two forms: the Overlay and the Perturbation methods (Knobs, 2001). These methods relate the level of the student knowledge with the learning objectives / competences that he intends to reach (Knobs, 2001).

2.3 Overlay

In this method, the user knowledge is related, layer to layer, to the Domain Model, producing the user knowledge model (Figure 3). The expression of the knowledge level of each concept is dependent on the Domain Model itself: this value can be binary (knows or ignores), qualitative (good, average, weak, etc.) or quantitative (the probability of knowing or not, a real value between 0 and 1, etc.).

In this method, the student knowledge is a subset of the system knowledge. The system does not allow representing the incorrect knowledge that the student acquired or might have acquired. This solution demands great flexibility in the student knowledge model for each topic (Brusilosky, 2001).

In addition, this method requires that the Domain Model represents individual topics and concepts. Its complexity depends on the granularity of the Domain Model structure and on the estimate of the student knowledge, acquired through the analysis of the student’s readings and assessments. This type of model is very flexible and capable of representing different Domains.

The techniques of Overlay and stereotype can be combined in Educational AHS. The student profile is initially categorized by one stereotype but is gradually modified when the Overlay Model receives information on the interaction with the system (Brusilosky, 2001).
2.4 Perturbation

This method considers that the knowledge and the student aptitudes are a perturbation of the specialist knowledge, and not a subset of his knowledge (as in the previous method) (Figure 4). This method can be used to represent knowledge that is beyond the Domain Model defined by the specialist.

![Figure 4. Representation of the Perturbation Model](image)

A method called Buggy Model, implemented in the program DEBUGGY (Brown, 1978), allows associating each rule to a set of wrong rules (Mal - rules). These rules are directly obtained from the tutor pedagogical experiences. The Student Model is obtained by replacing the correct rules with the wrong rules. When applied, they lead to the answers of the student. Since there can be several reasons for a student wrong answer (several wrong rules that lead to the student answer) the system proceeds to generate discriminating problems and presents them to the student to know exactly the wrong rules that this user has.

The second method consists in inferring the level of student understanding, using induction from the set of his answers (Laroussi, 2001). In this system, the student model has wrong knowledge but this information is obtained by the application of perturbation to the specialist knowledge.

3 Student Modeling Approaches

Different characteristics are used in some existing AHS (Table 2). Approaches to the UM in some of the AHS will be presented, next, systems that use overlay model will be described and then; the systems that use a combination of different techniques will be presented.
Some AHS that use the overlay model for UM are the following:

- The Adaptive Hypermedia Architecture (AHA) System is Educational AHS. The purpose of this system is to deliver courses over the web. The UM is based on concepts knowledge that the user acquires by solving tests and reading the hypermedia pages of the course.

- The AHM and XAHM system, in which the adaptation depends on the user’s level of expertise about the known concepts of the system domain (which is a subset of all domain concepts).

- The ISIS-TUTOR, which is a system, intended for learning the print formatting language of an information retrieval system CDS/ISIS/M which uses the overlay model with a set of integer counters.

- The HYPERFLEX, which is an adaptive hypertext browser. This system asks the user to specify his objectives and plans and uses a connected semantic network (Brusilovsky, 1996a; Brusilovsky, 1996b).

Many systems use stereotypes for describing the user. HYPERTUTOR is a system that only uses stereotypes for describing the user. This system employs exercises to obtain information about the users and uses stereotypes for UM. The student can belong to one of three groups: novice, medium or expert (Kavcic, 2000).

Many times one method alone does not allow the modeling needs of the system and the combination of diverse methods has to be chosen (Kavcic, 2000):

- ANATOM-TUTOR is a system to teach anatomy. It contains a rule-based user modeling component with operates with stereotypes and weighted rules.

- ELM-ART – Adaptive Remote Tutor is a system to support learning of Lisp programming language. It uses two UM Techniques: simple overlay model and complex Episodic Learner model. The user’s knowledge is represented by episodes that represent user individual learning history and his behavior and former problem solving situations.
• INTERBOOK is a tool for authoring and delivering adaptive electronic textbooks on the web. This AHS uses a concept based on the overlay model, but the UM is initialized using stereotypes.

• AVANTI is a system about metropolitan areas for a variety of users with different needs. This system combines an initial interview, stereotypes and the overlay method to create initial assumptions and then to maintain the knowledge of the user (Fink, 1996; Fink, 1997).

UM creation is also be achieved in:

• UMT (Brajnik, 1994): it allows the hierarchical definition of the user type through stereotypes, the definition of rules to infer the UM and the detection of contradictions. The user information received can be classified as invariable or assumptions.

• BGP-MS (Kobsa, 1995): it allows suppositions, represented by logical predicates, on the stereotype of the user or groups of users. Inference is achieved through different types of suppositions to define the user knowledge.

• DOPPELGÄNGER (Orwant, 1995): a server it accepts user information through hardware and software sensors. Several techniques to collect the sensor data are available. The users can visualize and edit their UM.

• TAGUS (Kobsa, 2001): it allows the definition of a stereotype hierarchy and contains an inference mechanism.

• UM (Kay, 1995): a UM toolkit, tries to represent suppositions of the user knowledge, preferences and others. The information is accompanied by a value that represents the confidence level.

4 Student Model Related Standards

The two main objectives of the standards are to allow interoperability and reuse of the learning tools and contents. The Learning Technology Standards Committee (LTSC) of IEEE or the Global Learning Consortium (IMS) are examples of organizations that define specifications and standards for e-learning.

The IMS Learner Information Package Information Model describes the structure of user data to allow interoperability between different systems. The Computer Society Standards Activity Board of LTSC has defined specifications like P1484 and P1484.2 for student data.

ISO/IEC JTC1/SC36 is formed by several Working Groups (WG). WG3 – Participant Information - intends to define the specifications related to the exchange of student information between different systems. It defines the data models, the syntax and the semantics to describe user characteristics, knowledge and abilities. It also describes the user acquisition of knowledge, capacities, aptitudes, personal information, relations, parameters of security, preferences and learning style, performance, portfolio, etc. In more detail:

• Biographical and demographic identification;

• Learning objectives and aspirations;

• Official qualifications, certifications and certificates;

• Activities related with education (formal or informal education);

• Work experiences;

• Trainings military or civic;

• Interests, occupations and activities of leisure;

• Abilities, aptitudes and knowledge;

• Cognitive and affective aptitudes and preferences;
• Psychomotor Domains;
• Accessibility, for the education information;
• Capacity of the language;
• Incapacities;
• Preferences of learning style;
• Physical preferences;
• Preferences techniques (for example operative system);
• Security Key and word key;
• Relations with other students.

The general information that must be stored is:
• Only one number of identification;
• First name and family name;
• Contact (cell phone or telephone and, optionally, email or/and address);
• Relations with the instructor and optionally the relations with the colleagues, the administrators of the course and/or with other structures;
• Security (word key, security key) and level of security.

The characteristics information to be stored:
• Learning objectives and options to prioritize each one;
• Diverse problems of the students (as preferred language, preferences techniques, qualifications, certificates, licenses, interests, affiliations, etc.);
• Cognitive characteristics (motivation, learning style, cognitive style, etc.);
• Tools of communications to accede the information of the course.

Table 3 represents the data and types considered relevant for the user identification. Table shows the field names and the definition of the date types.

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier</td>
<td>Octetstring</td>
<td>Only one student identification</td>
</tr>
<tr>
<td>Name_Family</td>
<td>String</td>
<td>Student family name</td>
</tr>
<tr>
<td>Name_list</td>
<td>String</td>
<td>Student first name</td>
</tr>
<tr>
<td>Telephone_hid_list</td>
<td>Digits</td>
<td>Telephones numbers associates to the student</td>
</tr>
<tr>
<td>Email_contact_list</td>
<td>String</td>
<td>Email associated with the student</td>
</tr>
<tr>
<td>Postal_address_list</td>
<td>String</td>
<td>Postal address associated the student</td>
</tr>
<tr>
<td>Full_name</td>
<td>String</td>
<td>Student full name</td>
</tr>
<tr>
<td>Sort_name</td>
<td>String</td>
<td>Last name, first name</td>
</tr>
<tr>
<td>DevicePreference</td>
<td>String</td>
<td>Designation preferred device name</td>
</tr>
<tr>
<td>Device_name</td>
<td>String</td>
<td>Preferred device name</td>
</tr>
<tr>
<td>Device_type</td>
<td>String</td>
<td>Description of the device preferred</td>
</tr>
<tr>
<td>Certificate_list</td>
<td>String</td>
<td>Certification associated with the student</td>
</tr>
<tr>
<td>Portfolio_hid_list</td>
<td>String</td>
<td>Student Portfolio</td>
</tr>
<tr>
<td>Organization_name</td>
<td>String</td>
<td>Name of the student organization</td>
</tr>
<tr>
<td>Organization_activity</td>
<td>String</td>
<td>organization Activity</td>
</tr>
<tr>
<td>TechnicalRequirements</td>
<td>String</td>
<td>Duration, size, requirements of the platform, etc</td>
</tr>
</tbody>
</table>
Standards are very relevant to allow communication and sharing of data, components and structures, at syntax and semantic level between systems. So, standardisation of the user model is an important issue, because through it is possible to improve the user model’s portability, as well as the interoperability of Educational System that apply standards for UM description. This will allow users to use different Educational Systems.

## 5 Conclusion

In the educational AHS, emphasis is put on the student knowledge related with the domain application, in the sense of making the most effective adaptation and allowing the student to reach his objectives (Chepegin, 2004).

The number and type of user characteristics to adapt depends heavily on the finality of each system, but some relevance is given to the cognitive part, learning styles and student knowledge (Brusilovsky, 2001; Brusilovsky, 2003; Chepegin, 2001; Of Bra, 2004a).

The analysis, application, implementation, integration and evaluation of techniques used to adapt the presentation and navigation in educational AHS, using metadata for the learning objects and user modeling, etc, will contribute to more value for implementation of e-learning in diverse academics institution, in a way to make the educational process more adaptive and capable to prepare future professionals.

At present, the research of the Student Model goes in the direction to make possible the reuse of a student model in different systems (Chepegin, 2004). The standards are more and more relevant for this effect, allowing the systems to communicate and to share data, components and structures, at syntax and semantic level (Chepegin, 2004), even if most of them still only allow syntax integration (Busilovsky, 2001; Of Bra, 2004a).
6 References


