Expert Systems with Applications 37 (2010) 2255–2264

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Acquisition, representation and management of user knowledge

Alejandro Peña Ayala *

WOLNM, 31 julio 1859, # 1099 B, Leyes Reforma, Iztapalapa, DF 09310, Mexico
ESIME, National Polytechnic Institute, Unidad Profesional "Adolfo Lopez Mateos", edificio 5, 3er piso, La Escalara, Gustavo A. Madero, DF 07738, Mexico
Centre for Computer Research, National Polytechnic Institute, Av. Juan de Dios Batiz s/n esq. Miguel Othon Mendizabal, Unidad Profesional "Adolfo Lopez Mateos", Zacatenco, Gustavo A. Madero, DF 07738, Mexico

ARTICLE INFO

Keywords:
Web-based Intelligent Systems (WBIS)
Web-based Education System (WBES)
User knowledge
Cognitive maps
Fuzzy-causal inference

ABSTRACT

Web-based Intelligent Systems (WBIS), e.g. information retrieval, intelligent Web, and e-Learning, deal with tasks such as acquisition, representation, and management of knowledge about users. Based on a user profile, WBIS are able to behave according the particular needs of people through the intelligent adaptation of services, content, navigation interfaces, and many more factors. Thereby, the design of an approach devoted to meet such tasks is critical for achieving the goals pursued by WBIS. Therefore, in this article an approach oriented to elicit, state, and administrate user knowledge is outlined. This work introduces a user model, which supports the selection of teaching experiences that are delivered to students in the e-Learning field. The aim is to enhance the apprenticeship of individuals that receive lectures according to the user model that a Web-based Education System (WBES) holds about them. According to a sort of empirical outcomes, it is concluded that: "The success of WBIS is biased by the accurately acquisition, representation, and management of user knowledge fulfilled by the approach".

1. Introduction

As a result of the Artificial Intelligence (AI) evolution and the spread of Internet, a mutual bias has grown from some of their trends. Actually, several AI scopes are enhanced when their specifications include the implementation of AI applications on the Web. Also, new Internet research lines have emerged in order to provide intelligent services, such as semantic Web and wisdom Web (Liu, 2003). While, Internet technologies tend to provide intelligent capabilities to a world community of users – which is increasing in number, demands, and services –, the AI research and applications pursue to spread their findings and outcomes world wide. Hence, WBIS have emerged like a kind of "joint venture" approach that takes into account paradigms from AI and Internet fields.

Before a plenty of users with different backgrounds, customs, and preferences, WBIS face up constraints and needs that traditional "one-size-fits-all" approaches are unable to satisfy (Brusilovsky, 1996). Thereby, WBIS require the support of user knowledge that reveals who is each user. Such knowledge is set and managed by a *user model*, which embraces a knowledge repository and an engine. A user model contains a sort of descriptions of what is considered relevant about people. It also provides advice to WBIS in order to make easy their adaptation to each individual (Koch, 2000). Thus, WBIS must include a user model with the purpose of customize their presentation, content, navigation, services, and interaction with users. Consequently, WBIS that take into account a user model, behave in an adaptive and intelligent way before people.

The cycle process for user modeling embraces three tasks: knowledge acquisition, knowledge representation, and knowledge management. This cycle is accomplished through several iterations, from a basic version to a suitable one that is appropriate for being used in real scenarios. Afterwards, the user model is continually updated as a result of future system maintenances.

In regards to the *knowledge acquisition*, this task is devoted to transfer knowledge from one or more sources to repositories managed by the engine of the user model. According to Chin (1993), the techniques for acquiring user knowledge are characterized along several orthogonal dimensions as follows: (1) *active* or *passive*, depending on the intensity of the user's participation during the acquisition; (2) *automatic* or *manual*, related to the degree of automation of the acquisition task fulfilled by the system; (3) *explicit* or *implicit*, according to the type of user feedback; (4) *online* or *offline*, based on when the acquisition is performed; (5) *adaptive* or *adaptable*, whether the system or the individual updates the user model respectively. In regards to machine learning for user modeling, Webb, Pazzani, and Billsus (2001) point out four issues to be considered: large data sets, labeled data, concept drift, and

* Address: WOLNM, 31 julio 1859, # 1099 B, Leyes Reforma, Iztapalapa, DF 09310, Mexico. Tel.: +52 55 5694 0916/5454 2611; fax: +52 55 5694 0916.
E-mail addresses: apenaa@ipn.mx, apenaa@wolnm.org, apenaa@sagitario.cic.ipn.mx.
URL: http://www.wolnm.org/apa.

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.07.047

computational complexity. Also, Zukerman and Albrecht (2001) focus on seven types of predictive statistical frameworks for user modeling: linear, term frequency inverse document frequency, Markov, neural networks, classification, rule induction, and Bayesian networks. Moreover, Chin (2001) points out several psychological tests for measuring people's aptitudes, cognitive skills, and personality, which are used for user modeling.

Regarding to the *knowledge representation*, this task deals with three items: *content*, *structure*, and *notation*. The content concerns with the nature of knowledge to be stated in the user model. Typically, the user model content points out the knowledge domain of the application, independent knowledge domain, and psychological properties, such as cognitive, learning preferences, goals, and personality of the individual (Koch, 2000). The structure corresponds to the organization style that is used to set the user model. According to Zaitseva et al. (2003), the most common structures for student modeling are the following: vector, net, scalar, genetic graph, and simulation. A vector structure is suitable for depicting psychological attributes. Net structures are appropriate for sketching knowledge domain. Scalar structures are devoted to identify users' achievements. A genetic graph describes the development of skills and the knowledge learned by the student. Simulation models point out student's knowledge domain acquired as a data structure and her/his skills as procedures, which are activated by an interpretative mechanism. In regards to the notation, the most common practice is the use of propositional calculus. For instance, the formal model for the user model, given by Self (1991), is set by formulae that depict the objects of beliefs that a system holds about the user. Another classical notation is a sort of concepts that are evaluated by Boolean, quantitative, qualitative or probabilistic values (De Bra & Calvi, 1998).

Finally, the *knowledge management* represents the third task in the user model's life-cycle. This task pursues to provide user knowledge that enables an adaptive behavior for the system. So according to the nature of WBIS a user model must provide suitable user knowledge. For instance, in the e-Learning field, a WBES aims to teach students as well as possible depending on a pedagogical paradigm. Thereby, the user model manages the knowledge according to such reference. When the WBES is concerned with the student's apprenticeship, an *overlay* user model manages the universe of knowledge acquired by the user as a subset of the knowledge domain that the expert holds (Slemman & Brown, 1982). But, when the pedagogic paradigm focuses on the student's behavior, the *perturbation* user model gives away the discrepancies between the user's behavior and the expert's one when they deal with a given problem. However, if the target of the WBES is the misconceptions of the student, a *bug* user model states bugs and mistakes that the user reveals during the interaction with the system (Vassileba, 1992).

Based on the sequence of tasks for building a user model, the organization of the paper is as follows: In section two the knowledge acquisition task is outlined. Essentially, the techniques applied for user modeling meet the active, automatic, explicit, online, and adaptive characteristics. The third section is devoted to represent the user model. In this approach the aim is to depict some psychological properties of the user and the knowledge domain acquired. The structure is a kind of net that reveals ontological classes. The notation is based on concepts, whose states are instantiated by fuzzy values. In section forth, the management of knowledge is done according a *predictive* user model. Such model takes into account the causality phenomenon. The fifth section is oriented to resume the experimental trial that was achieved to test the approach. In the conclusions section the main outcomes and the future work to be achieved are stated.

2. Knowledge acquisition

The user model is a kind of *mental model* that a system builds to gather relevant knowledge about an individual. On the other hand, the acquisition of user knowledge is a complex task that is the target of study of several disciplines such as psychology, sociology, and neural sciences. Study and practice of knowledge acquisition are rapidly gaining importance, because they are widely recognized today as a "bottle-neck" and the most critical part in the development of an AI system. Hence, in this section a method for acquiring and representing user knowledge is stated. Besides, the techniques used for knowledge acquisition are described. Moreover, some tools devoted to elicit information from users are introduced. Given these items, the Web-based knowledge acquisition approach is outlined next.

2.1. Methods devoted to the knowledge acquisition and the knowledge representation

According to the knowledge engineering, in this approach a method oriented to acquire and represent knowledge is proposed. This method embraces the following seven steps:

- *Identification* answers four basic questions: (1) Which is the system's goal to be achieved? (2) Which is the content that the user model should contain? (3) Which are the domains that the user model encompasses? (4) What are the criteria that user model must meet to become useful?
- *Conceptualization* uncovers primitive concepts and relationships between them.
- *Epistemological analysis* identifies structural properties of the conceptual knowledge, such as a taxonomy that is organized in terms of hierarchy, causal, spatiotemporal, or part-whole relationships.
- *Formalization* deals with issues about certainty, completeness, reliability, and consistency of knowledge.
- Logical analysis aims at knowledge about how to perform reasoning tasks in the domain.
- *Implementation* designs the structure and chooses the notation required for depicting knowledge.
- *Testing* validates the acquired knowledge that is stemmed from a large and representative sample of test cases.

2.2. Knowledge acquisition's techniques

First of all, the four questions stated in the *identification* step of the knowledge acquisition method are answered next based on their respective sequence: (1) the WBES pursues to enhance the apprenticeship of people based on the user model about them (Canales, Peña, Peredo, Sossa, & Gutiérrez, 2007); (2) psychological properties of the individual and knowledge domain that the user holds; (3) learning preferences, cognitive skills, and personality attributes about the individual; also, background and acquired knowledge domain of the student; (4) the apprenticeship of people who take lectures based on their user knowledge will be higher than the learning that their peers acquire from lectures that are randomly chosen without considering any user knowledge.

Afterwards, knowledge acquisition techniques are selected according to the orthogonal dimensions, previously stated, as follows: (1) *active*, the user is fully involved in the elicitation of knowledge about her/him through the application of psychological and knowledge domain tests; (2) *automatic*, based on psychological and pedagogical paradigms, the approach achieves a high degree of automation; (3) *implicit*, the feedback is stemmed from the