A Recommender System for Learning Goals

Nicola Capuano[1], Roberto Iannone[1], Matteo Gaeta[1,2], Sergio Miranda[2],
Pierluigi Ritrovato[2] and Saverio Salerno[1,2]

CRMPA, Centro di Ricerca in Matematica Pura ed Applicata,
Via Ponte don Melillo, 84084 Fisciano SA, Italy.
{capuano, iannone, gaeta, salerno}@crmpa.unisa.it

Dipartimento di Ingegneria Elettronica e Ingegneria Informatica, Università di Salerno,
Via Ponte don Melillo, 84084 Fisciano SA, Italy
{mgaeta, smiranda, pritrovato, salerno}@unisa.it

Abstract. The aim of a recommender system is to estimate the utility of a set of objects belonging to a given domain, starting from the information available about users and objects. Adaptive e-learning systems are able to automatically generate personalized learning experiences starting from a learner profile and a set of target learning goals. Starting form research results of these fields we defined a methodology to recommend learning goals and to generate learning experiences for learners of an adaptive e-learning system.

Keywords: e-learning, recommender systems, intelligent tutoring systems.

1 Introduction

A significant educational action able to guide the learner in a comprehensive learning process is not only focused on learning (cognition level) but also on fostering a correct learning behavior that empowers learners to achieve their learning goals in a controlled and directed way (metacognition level) [1].

Starting from this principle we defined and developed an e-learning system able to build personalized learning experiences starting from a set of target concepts selected on an ontology-based domain model [2]. We then extended such system in order to allow course generation form an explicit request in terms of needs to be satisfied and expressed by the learner in natural language [3].

The work presented in this paper deals with the definition of a further process of course building starting from an implicit request rather than from an explicit one. In other words, a methodology to recommend learning goals based on the analysis of a learner' profile (including known topics) and on the comparison of this profile with profiles of similar learners is defined.

The proposed methodology upholds the social presence while supporting the development of self-regulated learning. Educational recommendations serves as a pedagogical advance organizer for the learners, as it anticipates and spreads needs, knowledge and learning paths. Furthermore it also supports help seeking processes improving the students' control over learning.

The paper is organized in this way: the section 2 introduces some background about recommender systems and presents some existing application of such systems in e-learning; the section 3 briefly introduces the starting point of our research i.e. the learning system IWT; the section 4 describes the proposed methodology; eventually the section 5 describes conclusions and planned future work.

2 Background and Related Work

Recommender Systems (RS) are aimed at providing personalized recommendations on the utility of a set of objects belonging to a given domain, starting from the information available about users and objects.

A formal definition of the recommendation problem can be expressed in these terms [4]: C is the set of users of the system, I the set of objects that can be recommended, R a totally ordered set whose values represent the utility of an object for a user and u: $C \times I \rightarrow R$ a utility function that measures how a given object $i \in I$ is useful for a particular user $c \in C$. The purpose of the system is to recommend to each user c the object i that maximizes the utility function so that:

$$i'_c = \arg\max_{i \in I} u(c,i) . \quad (1)$$

The central problem of the recommendations is that the function u is not completely defined on the space $C \times I$ in fact, in typical applications of such systems, a user never expresses preferences on each object of the available catalog. A RS shall then be able to estimate the values of the utility function also in the space of data where it is not defined, extrapolating from the points of $C \times I$ where it is known.

Several approaches to recommendation exist in the literature. They are usually classified in three categories: *content-based approaches* recommend to a user objects similar to those that he have positively rated in the past; *collaborative approaches* recommend to a user those objects that are liked by other people with similar tastes; *hybrid approaches* combine the two previous approaches.

Several **RS for e-Learning** have been introduced to select and propose learning resources to users. One of the first system, based on a collaborative approach, has been Altered Vista [5]. Its goal was to explore how to collect user-made evaluations of learning resources and to propagate them in the form of recommendations about the qualities of the resources. A Similar system is RACOFI [6] that integrates a collaborative RS with a rule-based inference engine.

QSIA [7] is a RS for learning resources sharing, assessing and recommendation in online communities. CYCLADES [8] uses a collaborative approach with user-based ratings, but does applies the technique to several communities at the same time. A related system is CoFind [9]: it uses digital resources that are freely available and applies for the first time folksonomies for recommendations.

Shen and Shen [10] developed a recommender system for learning objects that is based on sequencing rules that help users be guided through the concepts of an ontology of topics. A similar sequencing system is LSRS [11] that analyzes grouplearning experiences to predict and provide a personal list for each learner by tracking others' learning patterns regarding certain topics.

In ReMashed [12] learners can rate information from an emerging amount of Web 2.0 information of a Learning Network and train a recommender system for their particular needs. The CourseRank system [13] uses instead a hybrid recommendation approach and is used as an unofficial course guide for Stanford University students. In the APOSDLE project [14] a contextual recommendations is offered to the employees of large organizations in the context of a knowledge-sharing environment.

3 The Starting Point

In this section we introduce a learning system named **IWT** (Intelligent Web Teacher) that we adopted as a basis to apply models and methodologies hereafter defined. As described in [2] IWT allows to generate personalized learning experiences and relies on four interacting models as described below.

The *domain model* describes the knowledge that is object of teaching through a set of concepts (representing topics to be taught) and a set of relations between concepts. A set of *teaching preferences* can be added to the domain model to define feasible teaching strategies that may be applied for each available concept.

The *learner model* represents a learner and is composed by a *cognitive state* that measures the knowledge reached by him at a given time and by a set of *learning preferences* that provide an evaluation of which learning strategies are more feasible for him. Both components are automatically assessed by IWT by analysing results of testing activities and the learner behaviour during the learning experience.

The *learning resource model* is a metadata representing a learning resource and is based on the application of the IEEE LOM standard [15]. It includes the set of concepts that are covered by the learning resource and an additional set of didactical properties representing learning strategies applied by the learning resource.

The *unit of learning model* represents a sequence of learning resources needed for a learner in order to understand a set of target concepts in a given domain.

In [2] we have described the process to generate a unit of learning starting from a set of a target concepts and from a learner model. The process generates a feasible sequence of domain concepts able to teach the target concepts. Then it removes domain concepts already known by the target learner by looking at his/her cognitive state. Eventually it associates to each remaining concept the best matching learning resources taking into account teaching and learning preferences.

To simplify user interactions with the system, IWT also implements an alternative method for the expression of a learning need through **Upper Level Learning Goals** (ULLG). An ULLG is a meaningful set of target concepts on a given domain model with a connected textual description [3]. ULLGs can be built either by teachers and by learners and are accessed through a search engine.

The learner can so specify a learning need in natural language and let the system find the list of best matching ULLGs basing on the similarity between the expressed need and the textual descriptions connected to ULLGs. Then the learner can select a

ULLG and let the system build a personalized unit of learning starting from the connected set of target concepts and from his/her learner model.

4 The Proposed Approach

This paper deals with the integration in IWT of a new process of course building based on ULLG but starting from an implicit request rather than from an explicit one. In other words, a methodology to recommend ULLGs based on the analysis of a learner' cognitive state and on the comparison of this cognitive state with cognitive states of similar learners is provided. In order to do so we will adapt and extend a user-to-user collaborative recommendation algorithm.

The algorithm consists of the following steps: concept mapping, concept utility estimation and ULLG utility estimation each described in one of the following a sub-sections. Once the utility of each ULLG is estimated for a learner, the ULLGs with the greater utility can be suggested to him.

4.1 Concept Mapping

Given a set of concepts C and a set of learners L, the *cognitive state* of a learner $l \in L$ (as reported in section 3 and detailed in [2]), describes the knowledge reached by l at a given time and it is represented as an application CS_l: $C \rightarrow [0, 10]$. Given a concept c, with $CS_l\,(c)$ we indicate the degree of knowledge (or grade) reached by the learner l for c. If such grade is greater then a threshold θ then c is considered as known by l, otherwise it is considered as unknown.

At a given time a learner can be enrolled to one or more units of learning. As reported in 3 (and detailed in [2]), a unit of learning represents a sequence of learning resources needed by a learner in order to understand a set of target concepts in a given domain. Among the components of a unit of learning there is the learning path $LPath = (c_1, \ldots, c_n)$: an ordered sequence of concepts that must be taught to a specific learner in order to let him/her complete the unit of learning.

Starting from that, we can define the set COT_l of all concepts that are object of teaching for a given learner as the union of all learning paths *LPath* corresponding to the units of learning the learner is enrolled in. Then we can define the **concept mapping function** that is a Boolean function CMF: $L \times C \rightarrow \{0, 1\}$ that can be defined as follows:

$$CMF(l,c) = \begin{cases} 1 \text{ if } CS_l(c) > \theta \text{ or } c \in COT_l \\ 0 \text{ otherwise} \end{cases}. \qquad (2)$$

So, given a leaner l, CMF(l,c) = 1 for all concepts c that are already known by l plus all concepts c that are currently object of teaching for him/her. It is equal to 0 for any other concepts.

4.2 Concept Utility Estimation

The utility $u(l,c)$ of a concept c for a learner l can be estimated starting from the concept mapping function. The utility of a known concept or of a concept that will be known soon is equal to 0. So $CMF(l,c) = 1 \rightarrow u(l,c) = 0$. Conversely, to estimate the utility of remaining concepts, a collaborative recommendation algorithm is used.

We can estimate the unknown utility of a given concept c for a learner l by aggregating, through a weighted sum, ratings for the concept c, included in the concept mapping function, coming for learners that are similar to l. The estimation can be done through the following formula:

$$u(l,c) = \frac{\sum_{l' \in L'} CMF(l',c) \cdot sim(l,l')}{\sum_{l' \in L'} |sim(l,l')|} \qquad \textbf{(3)}$$

where L' is the set of the n learners most similar to l while $sim(l,l')$ is the similarity degree between l and l' obtained though similarity measures like the *cosine similarity* or the *Pearson correlation coefficient* [4] calculated on *CMF*.

From the algorithmic point of view, to estimate the concept utility function, we start from the concept mapping matrix where each element $CMF(l,c)$ is defined with (2). This matrix is built the first time by considering every cognitive state and every course available on the system. Each time a learner starts, terminates or abandons a course then the row corresponding to this learner is updated, again, through (2).

Starting from the concept mapping matrix, the user-to-user similarity matrix is calculated. Each element $sim(l,l')$ of this matrix is obtained through a similarity measure between the rows of the concept mapping matrix corresponding to users l and l'. Once the similarity matrix is calculated, to estimate an undefined $u(l,c)$ for a given learner l, it is necessary to isolate and combine, by applying (3), the utility expressed for c by the n learners more similar to l.

4.3 ULLG Utility Estimation

An *ULLG* can be formally defined as a tuple $ULLG_i = (D_i, TC_{i1}, \ldots, TC_{in})$ where D_i is a text describing the learning objective in natural language, while $TC_1, \ldots, TC_n$ is the list of target concepts that have to be mastered by a learner in order to reach such learning objective. A learning need *LN* is a textual sentence (like “to learn Java programming” or “how to repair a bicycle” etc.) expressed by a learner in order to start the unit of learning building process.

Through the unit of learning generation algorithm introduced in section 3 (and detailed in [2]) IWT is able to generate a learning path starting from a set of target concepts. By applying the algorithm described there, it is possible to determine, for each existing upper level learning goal $ULLG_i$, the corresponding learning path $LPath_i$ starting from the connected list of target concepts.

Once determined learning paths associated to available ULLGs, it is possible to estimate the **aggregated utility** $au(l,ULLG_i)$ of each of them for a learner l with the following formula:

$$au(l,ULLG_i) = \sum_{c \in LPath_i} \frac{u(l,c)}{\left|LPath_i\right|} . \quad \textbf{(4)}$$

The calculus of the aggregated utility takes into account the utility of all concepts explained by the ULLG. This means that, if the learning path connected with the ULLG includes many concepts already known by the learner, its aggregate utility can be low even if the utility of remaining concepts is high. To take into account this information we introduce the concept of **marginal utility** $mu(l,ULLG_i)$ of $ULLG_i$ for a learner l that can be obtained with the following formula:

$$mu(l,ULLG_i) = \frac{\sum_{c \in LPath_i} u(l,c)\left(1 - CMF(l,c)\right)}{\sum_{c \in LPath_i} \left(1 - CMF(l,c)\right)} . \quad \textbf{(5)}$$

Thus the utility of an ULLG for a given learner can be obtained by combining aggregated and marginal utilities through a weighted sum with the following formula:

$$u(l,ULLG_i) = \alpha \; au(l,ULLG_i) + (1-\alpha) \; mu(l,ULLG_i) . \quad \textbf{(6)}$$

where α is the hybridization coefficient that is a real number between 0 (highest priority to the marginal utility) to 1 (highest priority to the aggregated utility). The choice for α will be done empirically basing on experimentation results. Low values for α privileges novelty while high values privilege accuracy of suggestions given by the recommender system.

5 Conclusions and Future Work

We defined in this paper a methodology to recommend learning goals and to generate learning experiences that will be integrated in IWT: an already existing adaptive e-learning system. The next step is to design and develop software components able to implement the defined methodology. An experimentation phase will follow to provide comments and suggestions to be used for models and methodologies improvement.

In addition to comments coming from experimentation, some improvement can be already foreseen. The application of *matrix factorisation techniques* [16] able to transform the concept mapping matrix that is an huge sparse matrix in a product of smaller dense matrixes can be applied to optimize recommender performances. In addition, the possibility for learners to rate ULLGs created by other teachers or learners will be explored. This rating can be exploited by recommender algorithms as explicit feedback to improve recommendations.

Acknowledgements. The research reported in this paper is partially supported by the European Commission under the Collaborative Project ALICE [17] “Adaptive Learning via an Intuitive, interactive, Collaborative, Emotional system”, VII Framework Program, Theme ICT-2009.4.2, Grant Agreement n. 257639.

References

1. Mangione, G. R., Gaeta, M., Orciuoli, F., Salerno, S.: A Semantic Metacognitive Learning Environment. In: Cognitive and Metacognitive Educational Systems, AAAI Fall Symposium (2010).
2. Capuano, N., Gaeta, M., Marengo, A., Miranda, S., Orciuoli, F., Ritrovato, P.: LIA: an Intelligent Advisor for e-Learning. In: Interactive Learning Environments, Taylor & Francis 17(3), 221-239 (2009).
3. Capuano, N., Gaeta, M., Orciuoli, F., Ritrovato, P.: On-Demand Construction of Personalized Learning Experiences Using Semantic Web and Web 2.0 Techniques. In: 9th IEEE International Conference on Advanced Learning Technologies, pp. 484-488 (2009).
4. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. In: Knowledge and Data Engineering, IEEE Transactions on Knowledge and Data Engineering 17(6), 734-749 (2005).
5. Recker, M. M., Wiley, D. A.: A non-authoritative educational metadata ontology for filtering and recommending learning objects. In Interactive learning environments 9(3), 255-271 (2001).
6. Anderson, M., Ball, M., Boley, H., Greene, S., Howse, N., Lemire, D., McGrath, S.: RACOFI: A Rule-Applying Collaborative Filtering System. In IEEE/WIC COLA 2003.
7. Rafaeli, S., Barak, M., Dan-Gur, Y., Toch, E.: QSIA-a Web-based environment for learning, assessing and knowledge sharing in communities. In Computers, Education 43(3), 273-289 (2004).
8. Avancini, H., Straccia, U.: User recommendation for collaborative and personalised digital archives. In International Journal of Web Based Communities 1(2), 163-175 (2005).
9. Dron, J., Mitchell, R., Boyne, C., Siviter, P.: CoFIND: steps towards a self-organising learning environment. In World Conference on the WWW and Internet (WebNet 2000), pp. 146-151, USA AACE (2000).
10. Shen, L., Shen, R.: Learning content recommendation service based-on simple sequencing specification. In: Liu W et al. (eds), LNCS, pp. 363-370, Springer (2004).
11. Huang, Y. M., Huang, T. C., Wang, K. T., Hwang, W. Y.: A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web. In Educational Technology & Society 12(2), 144-162 (2009).
12. Drachsler, H., Pecceu, D., Arts, T., Hutten, E., Rutledge, L., Van Rosmalen, P., Hummel, H. G. K., Koper, R.: ReMashed-Recommendations for Mash-Up Personal Learning Environments. In: 4th European Conference on Technology Enhanced Learning (EC-TEL 2009), LNCS 5794, pp. 788-793, Springer (2009).
13. Bercovitz, B., Kaliszan, F., Koutrika, G., Liou, H., Zadeh, Z. M., Garcia-Molina, H.: CourseRank: a social system for course planning. In 35th SIGMOD international conference on Management of data, ACM, (2009).
14. APOSDLE Advanced Process- Oriented Self- Directed Learning Environment. EU-project, http://www.aposdle.tugraz.at.
15. IMS Global Learning Consortium. IMS Meta-data Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata (2006).
16. Rendle, S., Schmidt-Thieme, L.: Online-updating regularized kernel matrix factorization models for large-scale recommender systems. In RecSys '08, the 2008 ACM conference on Recommender systems, ACM (2008).
17. ALICE Adaptive Learning via an Intuitive, interactive, Collaborative, Emotional system. EU-project, www.aliceproject.org.