ARTICLE IN PRESS
Journal of King Saud University – Computer and Information Sciences (2014) xxx, xxx–xxx

King Saud University
Journal of King Saud University – Computer and Information Sciences
www.ksu.edu.sa
www.sciencedirect.com

Using concept similarity in cross ontology for adaptive e-Learning systems

B. Saleena [a,*], S.K. Srivatsa [b]

[a] School of Computing Science and Engineering, VIT University, Chennai Campus, India
[b] Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai, India

Received 28 February 2013; revised 26 August 2013; accepted 13 March 2014

KEYWORDS
e-Learning;
Fuzzy domain ontology;
Cross ontology;
Semantic similarity measure

Abstract e-Learning is one of the most preferred media of learning by the learners. The learners search the web to gather knowledge about a particular topic from the information in the repositories. Retrieval of relevant materials from a domain can be easily implemented if the information is organized and related in some way. Ontologies are a key concept that helps us to relate information for providing the more relevant lessons to the learner. This paper proposes an adaptive e-Learning system, which generates a user specific e-Learning content by comparing the concepts with more than one system using similarity measures. A cross ontology measure is defined, which consists of fuzzy domain ontology as the primary ontology and the domain expert's ontology as the secondary ontology, for the comparison process. A personalized document is provided to the user with a user profile, which includes the data obtained from the processing of the proposed method under a User score, which is obtained through the user evaluation. The results of the proposed e-Learning system under the designed cross ontology similarity measure show a significant increase in performance and accuracy under different conditions. The assessment of the comparative analysis, showed the difference in performance of our proposed method over other methods. Based on the assessment results it is proved that the proposed approach is effective over other methods.

1. Introduction

The tutoring approach which nowadays is acquiring popularity all over the world, with progresses in Information and Communication Technology (ICT) is web-based education. Various organizations, institutes, universities, schools and corporations are spending considerable amounts of time and money in expanding online substitutes like e-Learning to conventional kinds of education and training systems in urbanized nations like the United States of America, the United Kingdom, and some European countries (Thyagharajan and Nayak, 2007; Ahmadpour and Mirdamadi, 2010). Such schemes should be proficient enough in delivering the appropriate content to a learner at the exact time in the most suitable way so as to offer customized instruction and must be capable of autonomously changing (update) its performance to assure the diverse requirements of learners (Jeon et al., 2007). Technical and domain based information are organized based

* Corresponding author.
E-mail address: saleenaameen@gmail.com (B. Saleena).
Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

1319-1578 © 2014 Production and hosting by Elsevier B.V. on behalf of King Saud University.
http://dx.doi.org/10.1016/j.jksuci.2014.03.007

on keywords and at present, they serve as the contents of e-Learning. Modern e-Learning should incorporate newfangled expertise in different learning activities, to accomplish extremely interactive as well as social-oriented education. However e-Learning cannot replace classroom guidance (Todorova, 2010; Šimún et al., 2007; Huang et al., 2006; Antonella and Carbonaro, 2008; Bhowmick et al., 2010).

In the e-Learning market the most proficient products are customized according to the requirements of the client (Alexakos et al., 2006). The capability to articulate World Wide Web information in a simple language that can be understood by machine and intelligent agents, thus allowing human users to trace, distribute and incorporate information in a mechanized manner was the aim of the Semantic Web (Chi, 2007). It provides a framework for dynamic, scattered and extensible structured knowledge (ontology) created on a formal logic. The semantics of the document is described in domain model (Mangalwede and Rao, 2010). Ontology is a formal representation of concepts and the relationship between them by means of an approved terminology endows with an affluent set of structures to put up a supplementary significant level of information. Ontology can get the straightforward appearance of a taxonomy of perceptions (i.e., lightweight ontology), or the further wide-ranging illustration of encompassing a taxonomy in addition to the axioms and constraints which distinguish several outstanding features of the real world (i.e., heavy weight ontology) (Bianchi et al., 2009).The semantic web technology has the potentiality to be employed in diverse areas. One of the domains which perhaps will take advantage from this web technology is e-Learning (Dutta, 2006). The ontology accounts for a common and instinctive way of the organization of a course (Colace et al., 2004). Concept maps are utilized to obtain and characterize the knowledge composition such as concepts and propositions as perceived by individuals (Horrocks et al., 2004). Concept maps are analogous to ontology in the sense that both of these tools are employed to correspond to concepts and the semantic relationships among concepts.

A number of methods have been developed to study and improvise the efficiency of the ontology based techniques for optimization of the process. Bianchi et al. (2009) introduced the use of Semantic Web services within Aqua Ring and ontology was used to support educational content explanation and retrieval. Snae and Brueckner (2007) presented an e-Learning management system with metadata that provided a general template for the situation of Thai learners. Ghaleb et al. (2006) put forward the research works in the field of e-Learning and also discussed the various applications of e-Learning, like virtual classrooms, distance learning and remote classrooms. Rene et al. (2011) proposed that ontology can be used in e-Learning to organize the teaching resources semantically by identifying the relationship between the materials, thereby improving the quality of teaching resources. Raymond et al., (2009) discussed that, with the extensive applications of electronic learning (e-Learning) technologies to education at all levels, an increasing number of online educational resources and messages were produced from the corresponding e-Learning environments.

The major differences between the proposed techniques with the existing techniques are discussed here. The work presented in Dutta (2006) provides the semantic web-based architecture for e-Learning. But, in the proposed technique, the retrieval of e-Learning contents is improved with the help of a semantic measure. In Snae and Brueckner (2007), the semantic e-Learning is developed for Thai learning environment. Here, the user profile-based learning environment is considered. The work presented in Ghaleb et al. (2006), Rene et al. (2011), Raymond et al. (2009), uses single ontology for providing the suitable contents or organizing the e-Learning contents. The proposed work aims in extracting resources from multiple ontologies using a semantic similarity measure thereby improving the retrieval of learning contents.

The paper is organized as follows: The second section provides models and expectation. The third section describes the methods utilized for the proposed e-Learning system. The fourth section deals with results and performance evaluation of the proposed approach under different criteria. Finally, fifth section discusses about the practical implication and sixth section concludes the paper with the scope for future enhancements.

2. Model and expectations

On considering the above discussed methods and their features, a new method is proposed to improvise the e-Learning system. The proposed approach is a concept similarity method, which is used to compare the similarities between the concepts in the different ontologies. The two different ontologies considered here are the fuzzy domain ontology and the Domain expert's ontology. The proposed approach deals with extracting details by comparing the concepts between the ontologies. The concepts are extracted from the concept map of the ontologies. A concept map is automated from the fuzzy domain ontology and a concept map is manually developed by experts with the help of domain expert's ontology. An XML file is generated based on a particular concept, which contains the representative and property set of the specific concepts from the ontologies. The representatives and properties are then processed for the generation of the concept similarity measure. The concept similarity measure is the main factor, which defines the most relevant concepts for the user according to the input query.

The main contributions of this paper are as follows,

- The use of more than one ontology, since most of the e-Learning system works with only a single ontology.
- The concept similarity measure, which we designed in the proposed method stands as the signature to the proposed method. The concept similarity measure is used to find related concepts from the different ontologies.
- The other feature in the proposed method is the user personalization for the e-Learning content. The model describes the user best knowledge and the information unknown to the user.

3. Methodology

3.1. Ontologies used in the proposed e-Learning system

In the current scenario, the ontology is considered as the hierarchy of concepts, which is a part of the concept map. The e-Learning systems (Zheng et al., 2013; Chikh, 2013;

Please cite this article in press as: Saleena, B., Srivatsa, S.K. Using concept similarity in cross ontology for adaptive e-Learning systems. Journal of King Saud University – Computer and Information Sciences (2014), http://dx.doi.org/10.1016/j.jksuci.2014.03.007

Hamani et al., 2013) that we consider are based on particular ontologies and each ontology is specialized in a particular concept map. In accordance with the latest researches, similarity between the concepts in the ontology can improve the performance of the e-Learning system. Keeping it on mind, we have planned to develop a method to illustrate the concept similarity in cross ontology. The cross ontology, deals with more than one ontology, so it also deals with similarities between the concepts in the different ontologies. The approach, that we have proposed considers two main ontologies,

(a) Fuzzy domain ontology.
(b) Domain expert ontology.

(a) Fuzzy domain ontology

This is an automated ontology, which is based on the technical domain that is used to develop the e-Learning content. This ontology serves as a stub for the e-Learning system (Raymond et al., 2009; Saleena and Srivatsa, 2014). The concept map is a taxonomical structure, which consists of a number of concepts that are interrelated to each other and the concepts are purely based on technical data. The concept map triggers the main function in the ontology for the smooth processing of the e-Learning content retrieval from the database. A sample ontology is illustrated in Fig. 1.

(b) Domain expert ontology

The domain expert ontology is the ontology developed by the experts for a particular domain. The expert ontology is very vast and effective ontology. The other important feature of the expert's ontology is that, the accuracy and precision are high. The expert's ontology has some limitations like size of the ontology and the execution time. The main disadvantage of domain expert's ontology with fuzzy domain ontology is the availability of experts with thorough knowledge on the domain which is required to construct the ontology.

The above illustrated ontologies are used for processing the proposed e-Learning system. The fuzzy domain ontology is basically constructed by collecting documents, which are specific to a domain or a topic. Thus, depending upon the information available and required by the ontology, we select a specific number of documents regarding the domain to construct the ontology.

3.2. Cross ontology similarity measure

The detailed description about the semantic similarity measure available in the literature for single and multiple ontologies is given in this sub-section. The proposed semantic cross ontology measure is also described using the proper mathematical equations.

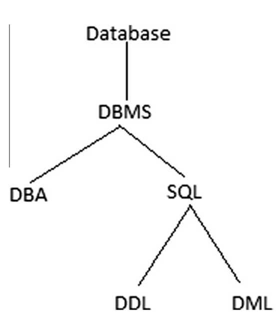


Figure 1 A sample fuzzy domain ontology.

3.2.1. Existing approaches on cross ontology similarity

Semantic Similarity (Hliaoutakis, 2005) computes the similarity between concepts (terms) which need not be lexically similar. An important consideration in finding the semantic similarity is whether it is carried out using a single ontology or multiple ontologies. When analyzing the literature, most of the works have concentrated only on incorporating single ontology to perform their similarity score. Accordingly, the detailed classification was given in Euripides et al. (2006), in which they classified the single ontology-based method into four main categories such as, edge counting methods, information content methods, feature-based methods and hybrid methods. All these methods are validated with two popular ontologies like WordNet or MeSH.

On the other hand, some of the researchers have used two ontologies to find the semantic similarity. Based on our knowledge, two popular similarity measures (cross ontology measure) (Euripides et al., 2006; Rodriguez and Egenhofer, 2003) were presented in the literature.

Rodriguez et al.'s measure: Rodriguez and Egenhofer (2003) have proposed a framework for comparing terms stemming from the same or from different ontologies. The similarity between terms a and b is computed as a weighted sum of similarities between synonym sets (synsets), features and terms neighborhoods:

$$Sim(a,b) = w.S_{synsets}(a,b) + u.S_{features}(a,b) + v.S_{neighborhoods}(a,b)$$

where w, u and v denote the relative importance of the three similarity components. The similarities among the synsets, features and terms neighborhoods are computed using the following equation.

$$S(a,b) = \frac{|A \cap B|}{|A \cap B| + \gamma(a,b)|A \setminus B| + (1-\gamma(a,b))|B \setminus A|}$$

where A and B denote synsets of terms a and b respectively and $A \setminus B$ denotes the set of terms in A but not in B (the reverse for B/A). Parameter $\gamma(a, b)$ is computed as a function of the depth of the terms a and b in their taxonomy:

$$\gamma(a,b) = \begin{cases} \frac{depth(a)}{depth(a)+depth(b)}, & depth(a) \leqslant depth(b); \\ \frac{depth(a)}{depth(a)+depth(b)}, & depth(a) > depth(b), \end{cases}$$

X-similarity: Followed by them, Euripides et al. (2006) have proposed X-similarity as cross ontology measure for computing the semantic similarity between terms stemming from different ontologies. In X-similarity, the similarity among the synsets and term description sets are computed using the following equation,

$$S(a,b) = \frac{|A \cap B|}{|A \cup B|}$$

where, A and B denote synsets or term description sets. Because not all terms in the neighborhood of a term are connected with the same relationship, they proposed another similarity formula for neighborhood sets.

$$S_{neighborhoods}(a,b) = \max \frac{|A_i \cap B_i|}{|A_i \cup B_i|}$$

where, i denotes relationship type (e.g., Is-A and Part-of for WordNet and only Is-A for MeSH). The above two ideas are combined into a single formula as follows,

$$Sim(a,b)=\begin{cases}1, & if\,S_{synsets}(a,b)>0;\\ \max\{S_{neighborhoods}(a,b),S_{descriptions}(a,b)\}, & if\,S_{synsets}(a,b)>0.\end{cases}$$

3.2.2. Designed cross ontology similarity measure for the proposed e-Learning system

In the proposed cross ontology similarity measure, the similarities of the concept in the two ontologies are measured on the basis of its neighborhood set and the feature set. We define a function for the similarity measure that incorporates the neighborhood and the features of the selected query keyword.

$$ConSim(c)=\sqrt{\frac{\alpha sim_{rep}^2+\beta sim_{pro}^2}{2}}$$

$$\alpha=\frac{O_{r_i}}{O_{r_i}+W_{r_i}}$$

$$\beta=\frac{O_{P_i}}{O_{P_i}+W_{P_i}}$$

where, $ConSim(c)$ is considered as the similarity measure between the concepts and α and β specify the relative importance of the two similarity constraints Sim_{rep} and Sim_{pro} respectively. Here $\boldsymbol{c}$ represents the selected concept, which contains the value according to the function of which it represents. O_r and O_P represent the number of representatives and properties in the primary ontology and W_r and W_P represent that of the secondary ontology.

The function Sim_{rep} is used for measuring the similarity of the representatives of the selected concept between the two ontologies. The similarity is measured by selecting both the representative sets, i.e. the one generated from the primary ontology and the one generated from the secondary ontology. The function listed below gives representative similarity values.

$$Sim_{rep}=\sqrt{\frac{R^2(O_i)+R^2(W_i)}{2}}$$

$$R(O_i)=\frac{(O_r^i\cap W_r^i)+\cup R(O)\cap\cup R(W)}{O_r^i},$$

$$R(W_i)=\frac{(O_r^i\cap W_r^i)+\cup R(O)\cap\cup R(W)}{W_r^i}$$

where $R(O_i)$ and $R(W_i)$ are the representative sets obtained from the primary and secondary ontologies respectively. The above function finds the relation between the matching neighbors and the non-matching neighbors of the concept, which is selected for the concept similarity. The relative value is a local maximum which is filtered under a specific threshold to get the most relevant and matching neighbors. The similar procedure is conducted again to get the features of the selected concept from the feature set, which we already acquired.

The value of Sim_{pro} for the selected concept and the elements in the feature set are again calculated, for finding similarity by comparing property sets of both the ontologies. The Sim_{pro} is also a local, maximum value. The property similarity function is given below,

$$Sim_{pro}=\sqrt{\frac{P^2(O_p)+P^2(W_p)}{2}}$$

$$P(O_p)=\frac{(O_p^i\cap W_p^i)+\cup P(O)\cap\cup P(W)}{O_p^i},$$

$$P(W_p)=\frac{(O_p^i\cap W_p^i)++\cup P(O)\cap\cup P(W)}{W_p^i}$$

where, $P(O_p)$ is the feature selected from the feature set of the primary ontology and $P(W_p)$ is the feature of the secondary ontology. These values are selected for the evaluation of the similarity of the selected concept through the modified equation from the Sim_{pro} equation.

The following summarize the differences between ConSim measure with the X-similarity (Euripides et al., 2006) and Rodriguez et al.'s measure (Rodriguez and Egenhofer, 2003):

- Rodriguez et al.'s measure has considered the depth of the terms in the two ontologies. However, cross ontology matching should not depend on ontology structure information (Euripides et al., 2006). The proposed measure and X-similarity have not considered the depth for computing similarity measure.
- Giving the appropriate weights will surely improve the similarity measurement. In Rodriguez and Egenhofer (2003), there are no appropriate weights given. On the other hand, X-similarity does not consider the weights in similarity computation. But, in the ConSim measure, we have provided the automatic weights for improving the similarity measurement.
- Both measures (Euripides et al., 2006) have not considered the *similar term* in the universal sets (synonyms, features and neighbors) of both the ontologies. Here, the similarity of universal set is also included so that a similar term presented in synonym sets and features set of both the ontologies is considered.

3.3. Adaptive e-Learning system using cross ontology similarity measure

e-Learning system is getting special attention in the field of online education. The purpose of e-Learning system is to provide an interface between student, instructor and knowledge database. There are a number of systems emerged to control the e-Learning systems, some of the latest approaches are discussed in the above section. In accordance with the review of the existing approaches, we have indented to propose a new approach for the e-Learning system. The proposed approach is an ontology based approach, which deals with concept maps. The proposed approach consists of a study, which illustrates the effect of concept similarity. In accordance with the latest researches, similarity between the concepts in the ontology can improve the performance of the e-Learning system. Keeping that in mind, we have intended to develop a method, which illustrates the similarity between concepts in different ontologies.

Block diagram:

The block diagram representation of the proposed e-Learning system is given in Fig. 2.

3.3.1. Concept extraction from the ontologies

In this phase, we extract the concepts, which are relevant to the user request, and process it accordingly. The concepts are extracted in accordance with the matching of a concept in the fuzzy domain ontology and the Domain Expert ontology. In our proposed approach, we consider the fuzzy domain ontology as *the primary ontology* and the domain expert's ontology as *the secondary ontology*.

(1) Neighborhood and feature sets

The proposed approach is a user interactive model, so the initial phase consists of the section of query keyword. The query keyword is the user input to our proposed approach. The keyword represents a concept, which is in relation with the primary ontology. The proposed method fetches the keywords as the input and compares it with both the ontologies. The comparison is done by searching keyword in the concept map of the fuzzy domain ontology and the domain expert ontology.

The processing of keywords, generates two set of keywords from each of the ontologies. The two sets are the basic measures, which are used for the similarity measure. The set of keywords are generated according to the neighborhood and the features of the input keyword that are present in the ontology.

Representatives: This is the set of keywords, which are present in the neighborhood of the query keyword. The representative set is defined as the set of keywords, which consists of the parent nodes that are associated with the keyword. In our proposed approach we define the branches of the super class according to a constant and specific iteration, in the sense that, the number of iterations does not affect the time requirement badly. The two different representative sets of each ontology are listed as separate entities.

$$S_R = \{(O_{r_i}, W_{r_i}), 1 \leqslant i \leqslant m$$

where S_R represents the set of representatives extracted from the ontology and O_{r_i} and W_{r_i} represent the representatives from given ontologies. In the proceeding sections, these entities are used for the similarity measure.

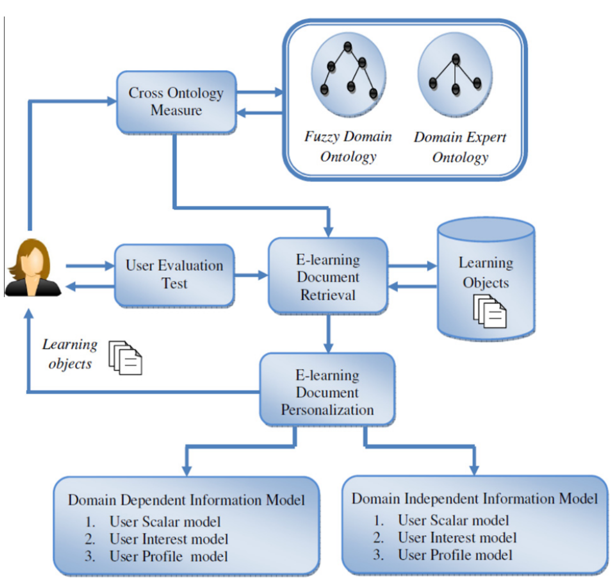


Figure 2 Block diagram of the proposed e-Learning system.

Properties: This set is a measure of part of relationship between the concepts in the ontology. The set consists of the keywords, which are relevant to the concept k. i.e., for the existence of concept k the keywords in the feature set are important. The feature set of k can be obtained with the help of the below stated function.

$$S_P = \{(O_{p_i}, W_{p_i}), 1 \leqslant i \leqslant m$$

The set S_P represents the properties selected from the given ontologies. The value of the function S_P is maximized to obtain the result regarding our proposed approach.

The above segment of ontology in Fig. 3 shows the representatives (super nodes) and properties (child nodes) for the concept *SQL*. The selected keywords are then listed as the features (properties) and neighbors (representatives) of the concept *SQL*. The representatives and properties of the query keyword are extracted from the XML description of the ontologies.

3.3.2. Concept similarity measure

The similarities of the concept in two ontologies are measured on the basis of its neighborhood set and the feature set. The selected concept is matched in the two ontologies according to the similarity of its neighbors and features in the primary ontology and the secondary ontology. The neighborhood set and feature set for the selected concept will be extracted for both approaches as per our proposed approach. The next phase deals with finding the similarity of the concept between the primary and secondary ontologies using the similarity measure given below.

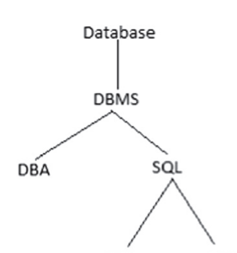
$$ConSim(c) = \sqrt{\frac{\alpha sim_{rep}^2 + \beta sim_{pro}^2}{2}}$$

The detailed description of designing the above similarity function is given in Section 3.2.2. The value of the function is used as the global maximum value, which is processed under a threshold to obtain the relevant concepts. The $ConSim(c)$ function is used as the main criteria for the similarity measure. The value obtained from the $ConSim(c)$ equations is utilized for the document retrieval process.

3.3.3. e-Learning document retrieval according to the concept similarity measure

The $Sim(a)$ functions generate a set of values according to the values generated from the functions Sim_{rep} and Sim_{pro}.

We provide a threshold for the values listed by the $ConSim(c)$ function and if the value is a higher value than the threshold we select that value, the rest of the values are neglected. Even though, the similarity measure gives enough



Node: *SQL*	
Super Node	Child Node
DBMS	DDL,DML
Database	DDL,DML

Figure 3 Representatives and properties of ontology.

Please cite this article in press as: Saleena, B., Srivatsa, S.K. Using concept similarity in cross ontology for adaptive e-Learning systems. Journal of King Saud University – Computer and Information Sciences (2014), http://dx.doi.org/10.1016/j.jksuci.2014.03.007

information for the retrieval of the document, to give more user specific results, we incorporate another constraint in the proposed approach. The constraint that we include deals with the user knowledge about the query, which he/she has given as the input. The constraint is formed by providing a set of questions to evaluate the user. The score obtained from the evaluation process is set as a conditional parameter. The user score and the value of $ConSim(c)$ are compared and in accordance with that, the document is retrieved to the user.

The above algorithm shown in Fig. 4 illustrates the process of document retrieval in accordance with the values generated from the concept similarity measure and based on the user score obtained from the user knowledge Information.

3.3.4. e-Learning document personalization based on user knowledge information

The user profile consists of the information about the user and the awareness of the user to the specific concept or the area of study. We provide a personalized system for the user by using the retrieved documents. The personalization is done according to a user profile, scalar and interest model which is illustrated in Fig. 5. The user profile model is designed in order to help the user by considering the user as a learner. A learner model must contain information about the learner's domain knowledge, his/her goals, interests, preferences and other information about the learner. Keeping the above mentioned information in mind, a user personalization model is designed, which is detailed below. The personalization is done based on the user score obtained from the user evaluation test. On the basis of the user score, we segregate the data information into two different models.

a. Domain dependant information model.
b. Domain independent information model.

(1) Domain dependent information model

Domain specific information represents a reflection of the learner's state and the level of knowledge and skills in a particular subject. In our proposed approach, the domain specific information of the user is measured from the user evaluation test. The user evaluation test is a sample test, which contains a number of queries related to the domain selected by the user. A user score is extracted from test and is used for the evaluation of the user. According to the user score, we deliver the learning contents to the user. The information we extract is based on the user score, which is delivered as positive responses from the user. According to the positive responses, we serve documents to the user based on three profiles.

(a) User scalar model.
(b) User interest model.
(c) User profile model.

Start
Select Sim_{pro}
Select Sim_{rep}
Compute $ConSim(c) = \sqrt{\dfrac{\alpha sim_{rep}^2 + \beta sim_{pro}^2}{2}}$
While ($ConSim(c)$ *>threshold***)**
  Select concepts (a)
End While loop
Select *User score*
Compare User score and $ConSim(c)$
Assign value to variable **C**
While (C)
Retrieve documents // documents from the database are retrieved
End

Figure 4 Pseudo code for document retrieval.

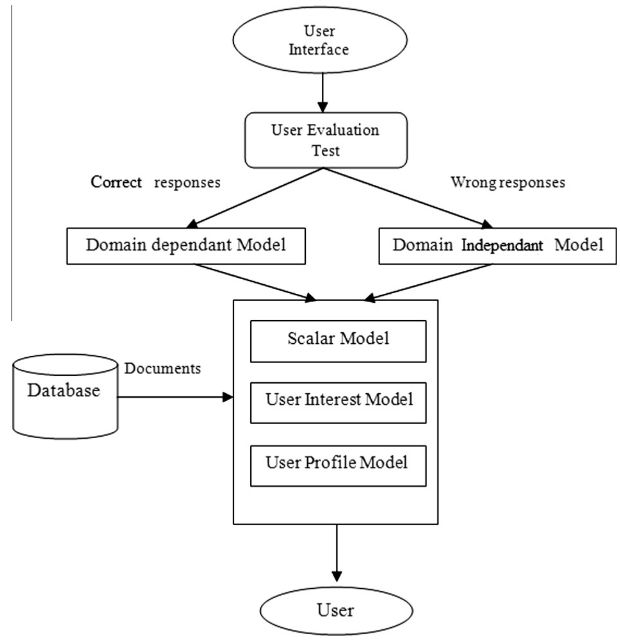


Figure 5 User personalization model.

We are using a scalar model for describing the information that is possessed by the user. In the scalar model, level of the learner's knowledge on the entire domain is described by one identifier such as a number in the range from 1 to 5. The scalar model is the simplest form of a knowledge model and provides no information about knowledge in a sub-domain (Fröschl, 2005). This section of the user model is modeled according to the scalar model as we discussed above, the identifier that we used in this context is the user score.

In the case of user interest model, we plot the data or document according to the interest of the user. The user may have specifications about the document such as, how the document should be. The user is interested in learning the content out of interest. Some users like to have diagrams and graphs in the learning content, some users like plan statements for learning. So according to the different scenarios we model the e-Learning Content.

The user profile model is a specification based on the designation of the user. The documents are provided to the user according to the degree level of the user. The criteria are based on the graduation level. If the user is a graduate, the document is provided according to extent of the information he/she needed and if a post graduate the information will change, as the level of the learner changes.

(2) Domain independent information

In addition to the learner's current knowledge level, for an adaptive e-Learning system, the domain independent information is also needed. This information contains different views of the learner to the current domain such as,

- *Goals*: To establish the correct teaching strategy, it is important to know the goals of learner like what he/she wants to achieve during e-Learning. The goals can be divided into two different types. First, the learning goal, which is relatively stable for a course unit. Second, the problem-solving goal, which may change from one problem to another even within one teaching unit.
- *Background and experience*: Background information includes skills that may affect the learning achievement. Such information is for example, profession, work experience or perspectives.
- *Preferences*: The learners may have different preferences related to some aspects of the learning environment. These preferences are considered as not inducible by the system. Thus, the learner has to inform the system directly or indirectly about his or her preferences. It is important for an adaptive e-Learning system to present and organize the learning material based on the learner's preferences. Learners can also be grouped based on their preferences.
- *Factual and historic data*: Demographic data such as name, age, parents, and ID. are often stored in learner models. This information, combined with other factual data such as for example interests, is necessary to initialize an individual learner model.

The user profile provides the full fledged idea of the leaner and the attitude of the leaner to the given problem. The proposed approach includes some extra features to the above proposed user profile model.

We use the user evaluation test for collecting the information about domain independent information also. The main difference between the domain dependant information and domain independent information is obtained based on the wrong responses from the user during the user evaluation test. In this section we define the user interest views and personal information, we also furnish the details that are needed for the user by evaluating the wrong responses. The domain independent information model is also based on the three profiles, scalar model, user interest model and the user profile model.

4. Experimental results

This section gives account of the analysis of the proposed approach under different test environments. The proposed approach is subjected to a test with a specific dataset to find the feasibility. The analysis is composed of two main analysis sections, performance analysis of the proposed method under different criteria and a comparison analysis with already existing MeSH ontology (Nelson et al., 2001) and WordNet Ontology (Ontolog). The section is detailed in the following section.

4.1. Experimental set up

The proposed algorithm is implemented using JAVA (JDK 1.7). For the experimental study, we collected e-Learning documents related with the domain "Database management System" from the web and the collected documents are arranged into 15 categories. For every category, one question is framed so that the 15 questions of 15 categories are used for user evaluation test. Furthermore, every category of documents is divided into two sub-categories that consist of documents with and without diagrams.

4.2. Experimentation of the proposed approach

(1) Two ontologies generated for experimentation

The proposed adaptive e-Learning system retrieves the e-Learning contents from the database and which is supplied to the learner. The processing of the adaptive e-Learning system is based on the ontologies we have discussed in the proposed approach. The two ontologies act as the center point of the cross ontology measurement. At first, the collected e-Learning documents are given to the previous approach (Saleena and Srivatsa, 2014) that generates the fuzzy domain ontology as shown in Fig. 6. An expert in the particular domain generates the domain expert ontology. The classes and sub classes are represented with different classes of id for their unique representation in the ontology. Analyzing the XML code generated, we can see the inter relationship between the class and sub classes. A more detailed representation of the ontology is shown in Fig. 7. It is the graphical representation of the ontology obtained through the fuzzy domain ontology. The XML code and the graphical representation of the ontology are generated using the software protégé 3.0 (Protege). It is a program used for generating OWL (Web Ontology Language). Similar to the fuzzy domain ontology, the protégé (Protege) is also used for the generation of XML code and graphical representation of the ontology for the domain expert's ontology. The XML code is represented in Fig. 8 and the graphical representation is shown in Fig. 9.

(2) Cross ontology measure for the different query words

The core idea behind our adaptive e-Learning system is the cross ontology measure of the concepts. This feature in our proposed method stood as the distinction over other methods. The cross ontology measure evaluates the relation between the concepts in a specific domain through the comparison of both the ontologies. The cross ontology specifies the class and subclasses according to the valued generated. In Table 1, we have plotted an example of the how the code are generated according to the cross ontology. From the XML codes, the elements represented in the tag Hypernyms are considered as the super classes and the tag Hyponyms are considered as the sub classes. The Hypernyms and Hyponyms of a concept from different ontologies are presented in Table 1.

As we mentioned above, whether a concept is super class or sub class of particular class is identified by the cross ontology measure between the concepts. Table 2 represents cross ontology measures calculated for different query keywords from the different domains.

The analysis of the cross ontology measure is detailed in the graph as shown in Fig. 10. The cross ontology measure stands for the inter relationship between the concepts between the ontologies. In the current scenario plotted in Table 2, we can see that, the concepts database_languages and database_types have a cross ontology measure of 2.056, i.e., they have a strong relation. In the similar way find out the concepts with higher inter relationship over the ontologies and hence considered for the E- learning content.

(3) e-Learning system

In an e-Learning system, the learning contents are termed as the most important of all features of the system. In our

Please cite this article in press as: Saleena, B., Srivatsa, S.K. Using concept similarity in cross ontology for adaptive e-Learning systems. Journal of King Saud University – Computer and Information Sciences (2014), http://dx.doi.org/10.1016/j.jksuci.2014.03.007

```xml
<?xml version="1.0"?>


<!DOCTYPE rdf:RDF [
    <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
    <!ENTITY swrl "http://www.w3.org/2003/11/swrl#" >
    <!ENTITY swrlb "http://www.w3.org/2003/11/swrlb#" >
    <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
    <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
    <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
    <!ENTITY protege "http://protege.stanford.edu/plugins/owl/protege#" >
    <!ENTITY xsp "http://www.owl-ontologies.com/2005/08/07/xsp.owl#" >
]>


<rdf:RDF xmlns="http://www.owl-ontologies.com/Ontology1318661397.owl#"
     xml:base="http://www.owl-ontologies.com/Ontology1318661397.owl"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
     xmlns:swrl="http://www.w3.org/2003/11/swrl#"
     xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
     xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:owl="http://www.w3.org/2002/07/owl#">
    <owl:Ontology rdf:about=""/>
    <owl:Class rdf:ID="Architecture">
        <rdfs:subClassOf rdf:resource="#Distributed_database"/>
    </owl:Class>
    <owl:Class rdf:ID="Cloud_database">
        <rdfs:subClassOf rdf:resource="#data_warehouse"/>
    </owl:Class>
    <owl:Class rdf:ID="data_warehouse">
        <rdfs:subClassOf rdf:resource="#Database"/>
    </owl:Class>
    <owl:Class rdf:ID="Database"/>
    <owl:Class rdf:ID="database_languages">
        <rdfs:subClassOf rdf:resource="#Sql"/>
    </owl:Class>
    <owl:Class rdf:ID="Database_models">
        <rdfs:subClassOf rdf:resource="#Database"/>
    </owl:Class>
```


Figure 6 XML file: fuzzy domain ontology.

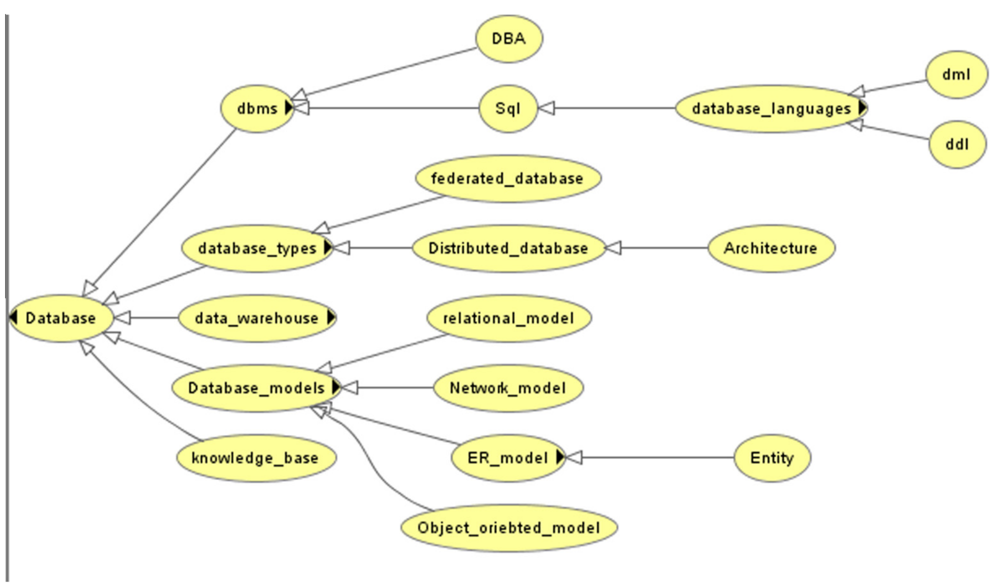


Figure 7 Graphic file: fuzzy domain ontology.

proposed approach, the retrieval is based on the user query and the user score from the user evaluation test. According to the user score, the proposed method defined a user profile model. The initial part of the model consists of the data, which are known to the user. These data are prepared according to the positive answers that are given by the user from the evaluation test.

The smoothness of the processing is based on the preparation of the documents, which are retrieved to the user. The documents are collected to a large database. The document is related to only a single domain as per the specification of user. The documents are retrieved to the user according to the specification that we discussed in the above sections. Three profiles are used for the process. We group the documents according to the different user models. When we consider the scalar model, there not much specification is needed, we simply plot the documents one by one. The user can select it as per his/her needs. In the case of user interest model, we processed the documents into two sections. The first section includes the documents with statements and diagrams and the second

```
</owl:Class>
<owl:Class rdf:ID="dbms">
    <rdfs:subClassOf rdf:resource="#Database"/>
</owl:Class>
<owl:Class rdf:ID="ddl">
    <rdfs:subClassOf rdf:resource="#database_languages"/>
</owl:Class>
<owl:Class rdf:ID="Distributed_database">
    <rdfs:subClassOf rdf:resource="#database_types"/>
</owl:Class>
<owl:Class rdf:ID="dml">
    <rdfs:subClassOf rdf:resource="#database_languages"/>
</owl:Class>
<owl:Class rdf:ID="Entity">
    <rdfs:subClassOf rdf:resource="#ER_model"/>
</owl:Class>
<owl:Class rdf:ID="ER_model">
    <rdfs:subClassOf rdf:resource="#Database_models"/>
</owl:Class>
<owl:Class rdf:ID="federated_database">
    <rdfs:subClassOf rdf:resource="#database_types"/>
</owl:Class>
<owl:Class rdf:ID="heirarchial_database">
    <rdfs:subClassOf rdf:resource="#Database_models"/>
</owl:Class>
<owl:Class rdf:ID="knowledge_base">
    <rdfs:subClassOf rdf:resource="#Database"/>
</owl:Class>
<owl:Class rdf:ID="Network_model">
    <rdfs:subClassOf rdf:resource="#Database_models"/>
</owl:Class>
<owl:Class rdf:ID="Object_oriebted_model">
    <rdfs:subClassOf rdf:resource="#Database_models"/>
</owl:Class>
<owl:Class rdf:ID="odbms">
    <rdfs:subClassOf rdf:resource="#types"/>
</owl:Class>
<owl:Class rdf:ID="oodbms">
    <rdfs:subClassOf rdf:resource="#types"/>
</owl:Class>
<owl:Class rdf:ID="oql">
    <rdfs:subClassOf rdf:resource="#database_languages"/>
```

Figure 8 XML code: domain expert's ontology.

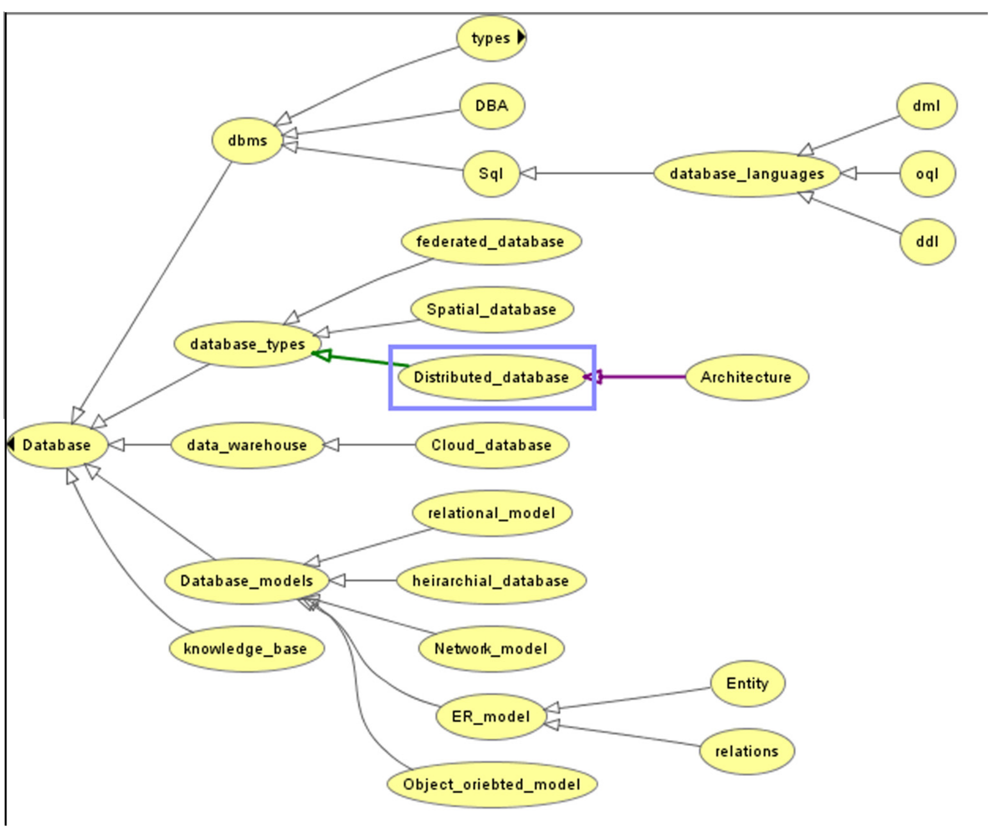


Figure 9 Graphic file-sample of domain expert's ontology.

Table 1 XML file generated for the two query keywords.

Fuzzy domain ontology	Domain expert's ontology
< Term > **database_languages** < Hypernyms > **sql**, < /Hypernym > < Hyponyms > **ddl**,**dml**,**dql**, < /Hyponyms > < /Term >	< Term > **database_type** < Hypernyms > **database**, < /Hypernyms > < Hyponyms > **active_database**,**analytical_database**, **cloud_database**,**data_warehouse**,**distributed_database**, **embedded_database**,**federated_database**,**real_time_database**,**spatial_database**,**temporal_database**, < /Hyponyms > < /Term >

Table 2 The cross ontology measure for different queries.

Q1 (Fuzzy domain ontology)	Q2 (Domain expert's ontology)	Cross ontology measure
Database	DBMS	1.0988
Database_languages	Database_types	2.056
Data_warehouse	Data_storage	1.803
Relational_model	Database_models	1.803

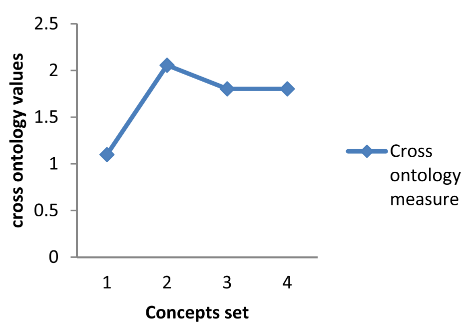


Figure 10 Cross ontology measure.

section deals with the documents and statements alone. The case is entirely different, when it comes with the user profile model. In dealing with the user profile model we have to group the document in different sections. The different sections are plotted on the basis of the degree or knowledge level of the user. According to these profiles, we retrieve the e-Learning content to the user.

4.3. Comparative evaluation with existing cross ontology measure and ontologies

The section, which we are going to furnish, deals with the comparative analysis of designed cross ontology measure over the two popular ontologies and two popular semantic similarity techniques. The ontologies, which we have taken into the account for comparison, are the WordNet (Ontolog) and MeSH (Nelson et al., 2001). The two cross ontology measures taken for comparison are the measures given in Euripides et al. (2006), Rodriguez and Egenhofer (2003). WordNet is a lexical database for the English language. It groups English words into sets of synonyms called synsets, provides short, general definitions, and records the various semantic relations between these synonym sets. MeSH (Medical Subject Headings) (Nelson et al., 2001) is a taxonomic hierarchy of medical and biological terms suggested by the US National Library of Medicine (NLM).

Here, the different query words are given to the previous similarity measure and the proposed similarity measure to find the semantic similarity. Table 3, describes the comparison of terms extracted from the two ontologies. The comparison is done based on the three algorithms, X-similarity measure (Euripides et al., 2006), Rodriguez and Egenhofer (2003) and the proposed similarity measure. From the analysis it is clear that the proposed approach produces better results for the same queries than the other methods. The cross ontology measure that we defined in our method differentiates the proposed method from the other methods by finding even the minute relationships between the concepts. The main difference between the proposed method and the other methods is the enhanced concept similarity measure.

Fig. 11 represents the responses of different similarity measures to the ontologies WordNet and MeSH. The terms

Table 3 Comparative analysis.

Query keyword		X-similarity measure (Euripides et al., 2006)	Rodriguez (Rodriguez and Egenhofer, 2003)	Proposed similarity measure
WordNet	MeSH			
Anemia	Appendicitis	0	0	0.2814
Dementia	Atopic Dermatitis	0	0	0.1346
Malaria	Bacterial Pneumonia	0.133	0	0.1807
Sarcoidosis	Tuberculosis	0.4062	0	0.4816
Carcinoma	Neoplasm	0.17	0.04	0.8951

Please cite this article in press as: Saleena, B., Srivatsa, S.K. Using concept similarity in cross ontology for adaptive e-Learning systems. Journal of King Saud University – Computer and Information Sciences (2014), http://dx.doi.org/10.1016/j.jksuci.2014.03.007

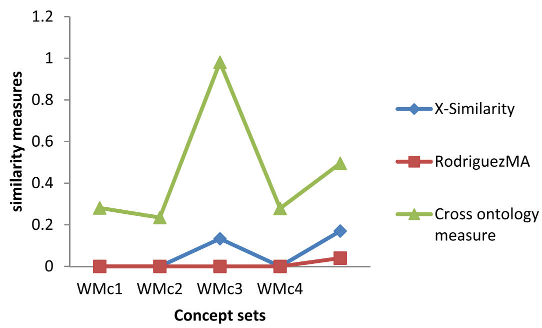


Figure 11 Comparative analysis.

WMc1and WMc2. represent the concepts from WordNet and MeSH.

5. Discussion and practical implication

The proposed approach is a concept similarity method, which uses two ontologies for the concept similarity. The approach we proposed is evolved from the two base ontologies, namely the fuzzy domain ontology and the wordnet ontology. The fuzzy domain ontology consists of the concepts and concept map, which are generated with the fuzzy concept and automatic concept map extraction algorithm. The fuzzy domain ontology contains the details about specific given data which may be technical or non-technical. On the other hand, wordnet is a commonly used ontology, which is a large lexical database of English, developed under the direction of George A. Miller. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept (Ontolog). Synsets are interlinked by means of conceptual-semantic and lexical relations. The proposed approach is based on extracting the concepts from these two ontologies for the concept similarity measure.

The fuzzy domain ontology is selected for our proposed approach because it possesses the details of a single domain only and more over it is used as concept map for the e-Learning system. The main advantages in selecting the fuzzy domain ontology are the semantic property of the ontology. The concepts in the fuzzy domain ontology have a strong semantic bond. The semantic bond helps the ontology to provide the concept based taxonomy as required by the user. As we discussed above, the wordnet ontology contains details about most of the domain, either technical or non-technical. The concepts maps are created according to data similarity algorithms. The proposed approach compares the concepts between the two ontologies to enhance the degree of relationship between a concept and its sub-concept in the concept map. The concepts with a higher degree of relationship are selected and given to the user.

The proposed approach defined a concept similarity measure over the concepts in both the ontologies. The concept similarity is based on two sets namely, representative set and property set. The representative set is extracted from the neighborhood of each concept. The property set is extracted by measuring the "IS-A" relation between the concepts. The proposed approach calculates the relation between the representatives from both the ontologies with the function Sim_{rep} and the relationship between the properties with the help of the function Sim_{pro}. The similarity measure is computed with the help of the above mentioned two functions. The similarity values obtained are filtered and then according to the similarity measure, we extracted documents from the database and supplied to the user through a user profile model. The user profile model is an interface given to the user, which contains the user's awareness about the current domain, the views and goals of the user and finally the user's relevant data to be studied or understand. The user profile model reduces the time and makes the e-Learning system more user interactive.

Future implications: The e-Learning systems are one of the most modern techniques in the education sector. The e-Learning provides a remedy for the difficulties faced in implementing fast learning system. So for the purpose of fast and effective learning, the E- Learning system should be an accurate and effective one. The proposed adaptive e-Learning system provides a different approach than the traditional e-Learning system. The cross ontology measure, which is used for the proposed approach is effective in finding relation between the concepts. The effectiveness of the proposed algorithm may help in providing better e-Learning solutions.

6. Conclusions and future enhancements

An adaptive e-Learning system with cross ontology similarity measure has been developed with automatically generated concept map as an application to the e-Learning system. The concept is designed with the most innovative and recent techniques. The cross ontology and the concept similarity measure have been incorporated for the novelty of the proposed method. The major feature in our proposed method is the user personalization model, which assesses the student capability of learning. The other attraction of our proposed method is that, we use multiple ontologies for the evolution of the concept from a particular domain. The use of more than one, ontology increased the performance and accuracy of the proposed method. The experimental result showed that our proposed system provides feasible results and it outperforms the existing methods. The futuristic enhancement to our system can be implemented using other data mining techniques and other data manipulation methods.

References

Ahmadpour, Amir, Mirdamadi, Mehdi, 2010. Determining challenges in the application of e-Learning in agricultural extension services in Iran. Am. Eur. J. Agric. Environ. Sci. 9 (3), 292–296.

Alexakos, Christos E., Giotopoulos, Konstantinos C., Thermogianni, Eleni J., Beligiannis, Grigorios N., Likothanassis, Spiridon D., 2006. Integrating e-Learning environments with computational intelligence assessment agents. Eng. Technol. 13, 117–122.

Antonella, Carbonaro, 2008. Ontology-based video retrieval in a semantic based learning environment. J. e-Learning Knowl. Soc. 4, 203–212.

Bhowmick, Plaban Kumar, Roy, Devshri, Sarkar, Sudeshna, Basu, Anupam, 2010. A framework for manual ontology engineering for management of learning material repository. Int. J. Comput. Sci. Appl. 7 (2), 30–51.

Bianchi, Stefano., Mastrodonato, Christian, Vercelli, Gianni, Vivanet, Giuliano., 2009. Use of ontologies to annotate and retrieve educational contents: the aquaring approach. J. e-Learning Knowl. Soc. 5 (1), 211–220.

Chi, Y.-L., 2007. Elicitation synergy of extracting conceptual tags and hierarchies in textual document. Expert Syst. Appl. 32 (2), 349–357.

Chikh, Azeddine, 2013. A general model of learning design objects. J. King Saud Univ. Comput. Inf. Sci..

Colace, Francesco, De Santo, Massimo, 2004. e-Learning: an assessment tool based on a Bayesian approach. Trans. Internet Res.,18–27.

Biswanath Dutta, 2006. Semantic Web Based e-Learning, DRTC Conference on ICT for Digital Learning Environment.

Euripides, G.M., Varelas, Giannis, Hliaoutakis, Angelos, Raftopoulou, Paraskevi, 2006. X-similarity computing semantic similarity between concepts from different ontologies. J. Digital Inf. Manage. 4.

Fröschl, C., 2005. User modeling and user profiling in adaptive e-Learning systems. Graz Austria Master Thesis.

Ghaleb, Fayed, Daoud, Sameh, Hasna, Ahmad, ALJa'am, Jihad M., El-Seoud, Samir A., El-Sofany, Hosam, 2006. e-Learning model based on semantic web technology. Int. J. Comput. Inf. Sci. 4, 63–71.

Hamani, Mohamed Said, Maamri, Ramdane, Kissoumc, Yacine, Sedratid, Maamar, 2013. Unexpected rules using a conceptual distance based on fuzzy ontology. J. King Saud Univ. Comput. Inf. Sci..

Angelos Hliaoutakis, 2005. Semantic similarity measures in MeSH ontology and their application to information retrieval on medline, Thesis Report.

Horrocks, I., Patel-Schneider, P.F., van Harmelen, F., 2004. From SHIQ and RDF to OWL: the making of a web ontology language. J. Web Semantics 1 (1).

Huang, Weihong, Webster, David, Wood, Dawn, Ishaya, Tanko, 2006. An intelligent semantic e-Learning framework using context-aware Semantic Web technologies. Br. J. Educ. Technol. 37 (3), 351–373.

Jeon, Sanghyun S, Su, Stanley Y.W. 2007. Achieving adaptive e-learning using condition-action rules and Bayesian networks. IADIS Int. Conf. e-Learning, 359–366.

Mangalwede, S.R., Rao, D.H., 2010. Application of bayesian networks for learner assessment in e-Learning systems. Int. J. Comput. Appl. 4 (4), 23–27.

Nelson, S.J., Johnston, D., Humphreys, B.L., 2001. Relationships in medical subject headings. Relation. Org. Knowl., 171–184.

WordNet ontology from <http://wordnet.princeton.edu>.

Protégé tool from "http://protege.stanford.edu/".

Lau Raymond, Y.K., Song, Dawei, Li, Yuefeng, 2009. Towards a fuzzy domain ontology extraction method for adaptive e-learning. IEEE Trans. Knowl. Data Eng. 21 (6), 800–813, June.

Rene, C.R., Robin, Uma, G.V., 2011. An ontology driven e-Learning agent for software risk management. Int. J. Acad. Res. 3 (2), 30–36.

Rodriguez, M.A., Egenhofer, M.J., 2003. Determining semantic similarity among entity classes from different ontologies. IEEE Trans. Knowl. Data Mining 15 (2), 442–456.

Saleena, B., Srivatsa, S.K., 2014. An Automated Approach to Extract Domain Ontology for E-Learning System, International Journal of Innovation and Learning(IJIL). Inderscience Publishers 15 (3), 254–273.

Šimún, Michal, Andrejko, Anton, Bieliková, Mária, 2007. Ontology-based models for personalized e-Learning environment. Comput. Inf. Sci., pp. 1–6.

Snae, Chakkrit, Brueckner, Michael, 2007. Ontology-driven e-Learning system based on roles and activities for Thai learning environment. Inter Dis. J. Knowl. Learning Objects 3, 1–17.

Thyagharajan, K.K., Nayak, Ratnamanjari, 2007. Adaptive content creation for personalized e-Learning using web services. J. Appl. Sci. Res. 3 (9), 828–836.

Todorova, Korneliya, 2010. Implementation of advanced technologies in ontology-based e-Learning model. Bulg. Open AIRE Rep., 260–264.

Zheng, Lanqin, Yang, Junfeng, Cheng, Wei, Huang, Ronghuai, 2013. Emerging approaches for supporting easy, engaged and effective collaborative learning. J. King Saud Univ. Comput. Inf. Sci..

Please cite this article in press as: Saleena, B., Srivatsa, S.K. Using concept similarity in cross ontology for adaptive e-Learning systems. Journal of King Saud University – Computer and Information Sciences (2014), http://dx.doi.org/10.1016/j.jksuci.2014.03.007