Journal of Theoretical and Applied Information Technology
20th December 2013. Vol. 58 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ADAPTIVE TEACHING AND LEARNING USING ONTOLOGY

[1] T. PAKIR MOITHEEN, [2] DR. P. SHEIK ABDUL KHADER

[1] Associate Professor, Department of Computer Applications, B.S. Abdur Rahman University, Chennai-48, India.
[2] Professor and Head, Department of Computer Applications, B.S. Abdur Rahman University, Chennai-48, India
E-mail: [1]pakirmoitheen@bsauniv.ac.in, [2]hodca@bsauniv.ac.in

ABSTRACT

The emergence of World Wide Web has provided more electronic resources to the learners in the e-learning field. The users are overloaded with electronic resources for a specific topic or domain. This research proposes a tool using ontology to evaluate the e-content and make the content adaptive to the need of the learner by generating concept maps for the electronic content. The tool also identifies the strength and weakness of the e-content by comparing with expert ontology. The tool is effectively used to personalize the content based on the knowledge level of the learner. The tool generated concept maps are 60% in accordance with expert ontology.

Keywords: *Adaptive learning, E-learning, ontology and personalization*

1. INTRODUCTION

Electronic learning (e-Learning) uses the information and communication technologies to enhance ordinary classroom teaching and learning. Now a day's users are overloaded with the e-learning resources and difficult to choose the best material for the specific topic. With the maturity of the internet technologies and the decreasing cost of the hardware platforms, more educational institutions are using e-Learning as a effective method for effective teaching learning process(1). The main objective of this project is to facilitate adaptive teaching learning so that instructors can dynamically revise and deliver instructional materials according to the learners' current progress. The current state-of the art of e-Learning technique uses automatic collection of learners' performance data using explicit test. However, few of the existing e-Learning technologies can support automatic analysis of learners' progress in terms of the knowledge structures they have acquired. In this paper, we illustrate a methodology of automatically constructing concept maps using ontology to characterize learners understanding for a particular topic, thereby instructors can conduct adaptive teaching and learning based on the learners knowledge structures as reflected in the ontology. In our approach we have used enhanced fuzzy domain ontology extraction algorithm. This paper focuses on developing a tool to evaluate the quality of eLearning content and gives best possible material to the student. It also overcomes the problem of information overloaded on the web by helping the student with this system. So this tool is useful for tutor to use adaptive teaching and for learners to improve their learning process.

Ontology is a formal specification of conceptualization (5). Ontology is the simple form of taxonomy of concepts (i.e., light weight ontology). Domain ontology is one kind of ontology which is used to represent the knowledge for a particular type of application domain. On the other hand, concept maps are used to elicit and represent the knowledge structure such as concepts and propositions as perceived by individuals (1). Concept maps are similar to ontology in the sense that both of these tools are used to represent concepts and the semantic relationships among concepts. However, ontology is a formal knowledge representation method to facilitate human and computer interactions and it can be expressed by using formal semantic markup languages such as RDF and OWL (6), whereas concept map is an informal tool for humans to specify semantic knowledge structure.

2. RELATED RESEARCH

Most of the e-Learning systems provide web-based learning so that students can access the same online courses via the internet without adaptation. In an e-Learning system, one size does not fit all. Therefore, it is a challenge to make e-Learning systems that are suitably "adaptive". The aim of adaptive e-Learning is to provide the students the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

appropriate content at the right time, means that the system is able to determine the knowledge level, keep track of usage, and arrange content automatically for each student for the best learning result(2).

Since now most of the e-learning systems have not been personalized completely. Several works have been carried out for personalizing the e-learning systems. Till now most of the current e-learning systems deliver the same content to the learner with different profile. A number of personalized systems have relied on explicit information given by a learner and have applied known methods and techniques of adapting the presentation and navigation. A web based intelligent tutoring system for teaching Java objects to students to overcome the difficulties they face in the programming. The basic idea of this system is a systematic introduction into the concept of Java objects. The system presents the topic of Java objects and administers automatically generated problems for the students to solve. The system is dynamically adapted at run time to the student's individual progress. The system provides explicit support for adaptive presentation constructs (3).

Today we are in an era where drastic advancements in networking and information technology are in action. The learning process has also taken these advancements, as a result of which e-learning came to the scene. Personalization in e-learning will improve the performance of the system. Recent researches are concentrating on providing adaptability to the learning management systems, depending upon the varying user needs and contexts. Adaptability can be provided at different levels .Providing an adaptive learning path according to the context of the learners' is an important issue. An optimal adaptive learning path will help the learners in reducing the cognitive overload and disorientation, and thereby improving the efficiency of the Learning Management System (4).

E-learning can be truly effective when it provides a learner centric adaptive learning experience. The success of any e-learning system depends on the retrieval of relevant learning materials according to the requirement of the learner. This leads to the development of the adaptive e-learning system to provide learning materials considering the requirements and understanding capability of the learner (5)

We employ a hybrid lexico-syntactic and statistical learning method rather than a computationally expensive graph-based approach for ontology extraction (6). Moreover, we employ the notion of fuzzy ontology rather than crisp ontology to explicitly model the uncertainty arising in automated ontology extraction. There was also research work exploring the ideas of automatically extracting ontologies from teaching documents although the algorithmic details were not illustrated (7). Previous work had also employed the Term Frequency Inverse Document Frequency (TFIDF) heuristic developed from the field of IR to extract prominent concepts from electronic messages generated in e-Learning. A knowledge density score was developed based on the TFIDF term weighting formula to assess the extent of contribution to online knowledge sharing by individuals. Our document parsing approach also employs TFIDF and other linguistic pattern recognition method to extract concepts from text (1).

3. FRAMEWORK FOR ONTOLOGY CONSTRUCTION

The main challenge of automatic ontology extraction from textual databases is the removal of noisy concepts and relations (6). Based on this issue, our domain ontology extraction methodology in general and process in particular are designed to effectively filter the non-relevant concepts and concept relations from the concept space. Figure 1 depicts the proposed methodology of automatically generating ontology for a collection of educational resources. At the document parsing stage, our document parser will scan each message to analyze the lexico-syntactic elements available in the e-content. Stop words such as "a, an, the" are removed from the content since these words appear in any contexts and they cannot provide useful information to describe a domain concept. For our implementation, a stop word file is constructed based on the standard stop word file used in the SMART retrieval system.

Lexical patterns are identified by applying Part-of-Speech (POS) tagging to the source documents. We develop our POS tagger based on the WordNet lexicon and the publicly available API (http://wordnet.princeton.edu/). We treat each named-entity as a noun for subsequent linguistic pattern mining. After the tagging process, each token is stemmed according to the Porter stemming algorithm. During the concept extraction stage linguistic patterns are ignored to reduce the generation of noisy concepts.

Journal of Theoretical and Applied Information Technology
20th December 2013. Vol. 58 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Our text mining program focus on certain linguistic patterns such as “Noun Noun”, “Adjective Noun”, “Verb Noun”, etc.., and find the term association information and collecting the statistical data for those patterns only. This will reduce the generation of noisy concepts but also improve the computational efficiency of our ontology extraction process.

A text windowing process will be conducted by scanning adjacent tokens within a pre-defined window size of 10 words from left to right over all the documents. At the end of the windowing process, an information theoretic measure is applied to compute the co-occurrence statistics between the targeting linguistic patterns and other tokens appearing in the same text window across the corpus. To produce accurate concept representations, a dimensionality reduction method is applied to the filtered concept space to minimize the terms (features) used to characterize the concepts based on the principle of minimal information loss. After concept space reduction, the subsumption relationships among the domain concepts are computed according to our enhanced fuzzy relation membership function. A taxonomy of domain ontology is constructed according to our fuzzy domain ontology extraction algorithm and displays them on our Web-based e-Learning platform.

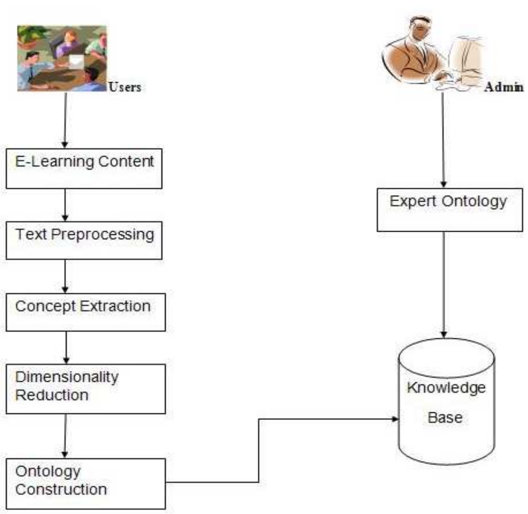

Figure 1 : Framework for Ontology Construction

The system provides two types of e-content evaluation, one is resource evaluation and the other is Students performance evaluation. The resource evaluation is to evaluate eLearning resource. The student performance evaluation is learners understanding for particular topic. So this algorithm facilitates adaptive teaching and learning process. The system architecture diagram for this adaptive e-learning system is depicted in Figure 1.

Algorithm

Automatic Enhanced Fuzzy Ontology Extraction.

Input: E- Learning Resources

Output: Ontology

Step 1: Ont = {}

Step 2: For each document Do

a) Construct text window of size w

b) Remove stop words from w

c) Perform POS tagging for each term ti in w

d) Apply Porter stemming to each term ti

e) Filter specific linguistic patterns

f) Accumulate the frequency for ti in w and the joint frequency for any pair ti, tj in w

Step 3: Perform Dimensionality Reduction SVD

Step 4: For each pair of concepts Ci, Cj Do

a) Compute the taxonomy relation r(Ci, Cj) using Specification(Ci, Cj)

b) IF M[RCC (Ci, Cj)] > μ, RCC = RCC U r(Ci, Cj)

Step 5: Output Ontology

Figure 2. The Automatic Fuzzy Domain Ontology Extraction Algorithm

4. ADAPTIVE E-LEARNING TOOL

The proposed system categorized into four phases Text Pre-processing, Concept Extraction, Dimensionality Reduction, Ontology Construction and Fuzzy taxonomy extraction.

A. Text Pre-processing.

Our text mining method is designed specially to filter the noisy concepts. In this text pre-processing stage, the following operations are carried out such as stop word removal, POS tagging, and word stemming. Stop words are removed from the documents such as “a, an, the” from the eLearning content as per the standard stop word file used in the SMART retrieval system. Lexical patterns are identified by applying Part-of-Speech (POS) tagging to the source documents using WordNet lexicon and the publicly available API (http://wordnet.princeton.edu/). After the

Journal of Theoretical and Applied Information Technology
20th December 2013. Vol. 58 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

tagging process, each token is stemmed according to the Porter stemming algorithm.

B. Concept Extraction

To extract concepts from the text corpora, a windowing process is conducted over the collection of documents. The windowing process can help reduce the number of noisy terms. For each document, a virtual window of 10 words is moved from left to right one word at a time until the end of a textual unit (e.g., a sentence) is reached. Within each window, the statistical information among tokens is collected to develop collocation expressions. To improve computational efficiency and filter noisy concepts, only the specific linguistic patterns (e.g., Noun Noun, Adjective Noun, etc.) will be analyzed. After parsing the whole corpus, the statistical data (e.g., mutual information) about the potential concepts is collected by our statistical token analyzer. If the association weight between a concept and a term is below a pre-defined threshold value, it will be discarded for ontology extraction.

$$MI(ti, tj) = \log 2\ Pr(ti, tj) / Pr(ti)Pr(tj) \quad \text{-------- (1)}$$

where MI(ti, tj) is the mutual information between term ti and term tj . Pr(ti, tj) is the joint probability that both terms appear in a text window, and Pr(ti) is the probability that a term ti appears in a text window. The probability Pr(ti) is estimated based on |wt| |w| where |wt| is the number of windows containing the term t and |w| is the total number of windows constructed from a corpus. Similarly, Pr(ti, tj) is the fraction of the number of windows containing both terms out of the total number of windows.

C. Dimensionality Reduction.

To produce accurate concept representations, a dimensionality reduction method is applied to minimize the terms (features) used to characterize the concepts based on the principle of minimal information loss. To reduce concept space, Singular value decomposition is applied.

D. Ontology Construction.

The final stage towards our ontology extraction method is fuzzy taxonomy generation based on the subsumption relations among the extracted concepts. Let Specification (Cx, Cy) denotes that concept Cx is a specialization (sub-class) of another concept Cy. The degree of such a specialization relation can be estimated from the following equation.

$$M(Cx, Cy) = \text{Specification}(Cx, Cy)$$

$$= M(Cx) \copyright M(Cy) / M(Cx) \quad \text{----- (2)}$$

Where M is fuzzy relation membership function and © is a fuzzy conjunction operator which is equivalent to the min function. The above formula states that the degree of subsumption (specificity) of Cx to Cy is based on the ratio of the sum of the minimal membership values of the common terms belonging to both concepts to the sum of the membership values of terms in the concept Cx. For instance, if every attribute of Cy is also an attribute of Cx, a strong specificity relation exists and the value of Specification(Cx, Cy) is high. The domain of the Specification(Cx, Cy) falls in the unit interval [0, 1] and the subsumption relation is asymmetric. After concept space reduction, the subsumption relationships among the domain concepts are computed according to our fuzzy relation membership function. A taxonomy of fuzzy domain concepts is then constructed according to our enhanced fuzzy domain ontology extraction algorithm.

E. Fuzzy Taxonomy Extraction

The fuzzy taxonomy extraction module is to extract the relation between the extracted concepts This module consists of two steps. The first step is to extract the relations between the concepts such that

$Spec(Cx, Cy) \geq Spec(Cy, Cx)$ and

$Spec(Cx, Cy) > \mu$ where μ is a threshold to distinguish significant subsumption relations. The parameter μ is estimated based on empirical tests.

If $Spec(Cx, Cy) = Spec(Cy, Cx)$ and

$Spec(Cx, Cy) > \mu$ is established, the equivalent relation between Cx and Cy will be extracted. The second step is pruning step, to remove the redundant taxonomy relations.

5. IMPLEMENTATION

The system has been implemented using java. The Stream Tokenizer class takes an input stream and parses it into "tokens", allowing the tokens to be read one at a time. The parsing process is controlled by a table and a number of flags that can be set to various states. The stream tokenizer can recognize identifiers, numbers, quoted strings, and various comment styles. String tokenizer is used in stopword removal module to parse the documents. JUnit is an open source framework designed for the purpose of writing and running tests in the Java programming language. JUnit, originally written by Erich Gamma and Kent Beck, has been important in the evolution of test-driven development, which is part of a larger

Journal of Theoretical and Applied Information Technology
20th December 2013. Vol. 58 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

software design paradigm known as Extreme Programming (XP). Text Pre-processing is the first module in this project, this module remove the stop words from the documents. Figure 3 shows the screen shots of the input documents.

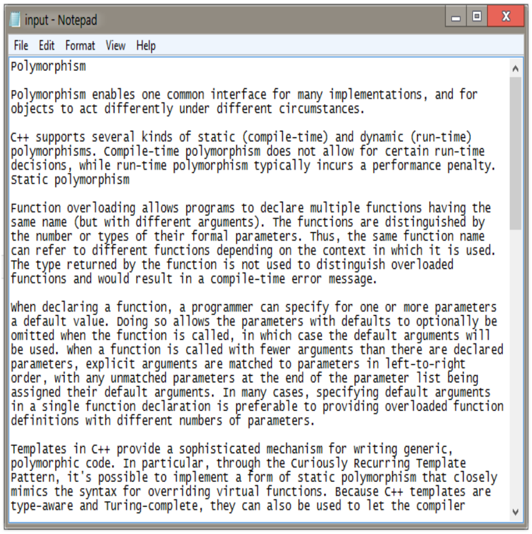

Figure 3 : Text Corpora

Next process in Text Pre-processing is stemming and POS tagging of the source document. Stemming of keywords involves extracting the root words such as plurals and gerunds. Some examples would include: running/run, apples/apple, and educational/education. In concept extraction module the important concepts in the document are extracted by applying windowing process. The windowing process scans the e-content from left to right by dividing documents by 10words per window until the end of the document is reached. It also removes duplicate word from the documents. Figure.4 shows the screen shot of concept extraction module. The System has generated concept map for java domain (polymorphism) using ontology is as shown in figure 5. Since the objective of the system is to analyze any eLearning document and provide student with the best study material. The created ontology is compared with the existing expert ontology of java programming language to check the eLearning document against expert ontology. Figure 6 shows the home page of Our system. The results obtained by the system compared with the expert ontology are 60%. So the tool can be used for personalizing the e-content by identifying the weakness in any java e-content.

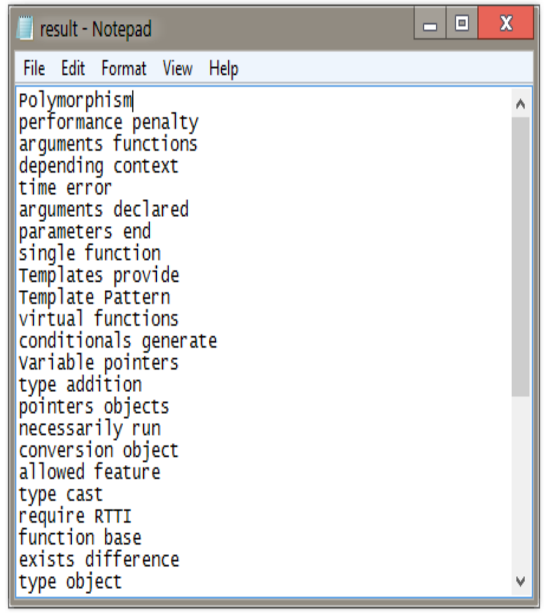

Figure 4: Concept Generation

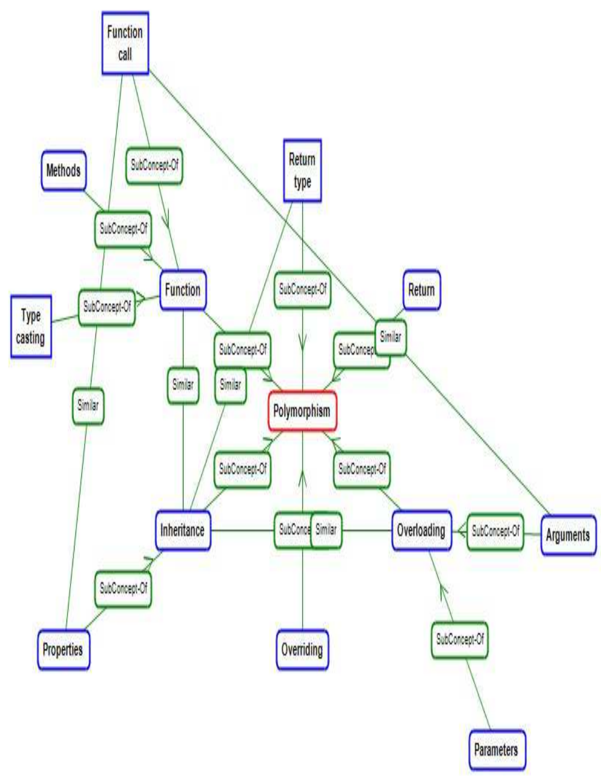

Figure 5: Ontology for Java Domain

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 6: Home page for e-learning system

6. CONCLUSION

In this research paper an e-learning tool is developed in order to help both the learners and the tutors to use the system effectively for teaching learning process. The system is working well and it is used by university teachers and students of B.S. Abdur Rahman Univerity. The new tool developed for adaptive learning purpose is well received by the student communities and faculties of our university. The results are promising and the concept maps generated by this tool are 60% in accordance with expert ontology. Since the tool needs expert ontology to evaluate the e-learning resources, we need subject experts to develop expert ontology for specific domain and topic. In future, it will be developed for analytical courses.

REFRENCES:

[1] R.Y.K Lau et al.., "Towards Fuzzy Domain Ontology Extraction Method for adaptive e-Learning", IEEE Transactions on knowledge and Data Engineering vol 2, No 6, June 2009.

[2] Esichaikul. V, Lamnoi, S, & Bechter, C..,"Student Modeling in Adaptive E-Learning System", Knowledge Management & E-Learning: An International Journal, Vol.3, No.3, 2011.

[3] Samy S. Abu Naser, ‘’ Evaluating the Effectiveness of the CPP-Tutor, an Intelligent Tutoring System for Students Learning to Program in C++’’, Journal of Applied Sciences Research, 5(1): 109-114, 2009.

[4] Pushpa. M., "ACO in e-Learning: Towards an adaptive learning path", International Journal on Computer Science and Engineering, Vol. 4 No. 03 March 2012.

[5] Subrat Roy, Devshri Roy., "Adaptive E-learning System: A Review", International Journal of Computer Trends and Technology-March to April Issue 2011

[6] Quan Thanh Tho, Siu Cheung Hui, Alvis Cheuk M. Fong, and Tru Hoang Cao. Automatic fuzzy ontology generation for semantic web. IEEE Transactions on Knowledge and Data Engineering, 18(6):842– 856, 2006.

[7] Papatheodorou, C., Vassiliou, A., Simon, B.: Discovery of ontologies for learning resources using word-based clustering. In: Proceedings of the World Conference on Educational Multimedia, Hypermedia and Tele communications (ED-MEDIA 2002), pp. 324–326, 2002.

[8] Zadeh, L.A.: Fuzzy sets. Journal of Information and Control 8, 338–353, 1965.