Effective Method for Automatic Generation of Knowledge from Electronic Textbook

Rasika A. Rode[1], Gitanjali Shinde[2]

[1] Dept. of Computer Engineering, Sinhgad Technical Education Society's, Smt. Kashibai Navale College Of Engineering, Pune, Maharashtra, India

[2] Assistant Professor, Dept. of Computer Engineering, Sinhgad Technical Education Society's, Smt. Kashibai Navale College Of Engineering, Pune, Maharashtra, India

Abstract: *With the tremendous increase in traffic and tools in virtually every organization we need such kind of tool that will perform all the operation regarding education system in effective way so that technology supported learning system should be helpful in many learning situations. we need to generate domain module and domain module contain the knowledge but generation of domain module require more cost and hence electronic textbook has been developed which is a core representation of the domain module. So, we use natural processing techniques, heuristic reasoning and ontology's for semiautomatic construction of domain module from electronic textbook.*

Keywords: Ontology design, Domain engineering, knowledge acquisition, semantic engineering, process and resource.

1. Introduction

In order to enhance both teaching and learning processes information and communication technology has affected education and technology supported learning system includes hypermedia system, intelligent tutoring system are being widely used in many academic institution in order to observed positive relationship between based technology and student engagement to web based technology and desirable learning outcome has been observed.

Since, building the knowledge is a hard task which not only involves the object to be learned but also require to identify the pedagogical relationship among those topics and this was the problem faced by textbook authors while they write their document and structured it for learning and hence, electronic textbook is considers as source to build the knowledge which helps in how the normal teacher prepares for the subject for this he considers set of references for getting knowledge like definition examples and exercises and schedule the lecture. Semiautomatic approach is considered so as to reduce development cost of knowledge.

Hence, to keep the updated ontologies with advance in knowledge there must be a method which will help in building such updated ontologies. So, whatever manual method was existing not was for improvement of scalability which was effort and time consuming in order to reduce effect that we observed previously automated method to build domain ontologies is necessary and hence we are building and evaluating knowledge generated from electronic textbook.

Semiautomatic method cannot be considers as holy grail which means can't be give clear and perfect structure. hence, we need knowledge generation techniques which includes ontology skeleton and all complex steps require for the generation of ontology skeletons. Another issue was there was lack of methodologies even they constructed

automatically or manually. But because of wide adoption of domain it provide and evaluate quality, richness, cohesion and domain covering.

The one important feature provides generation of attributes and relationship between them which are richly described and density measures uses in overall evaluation the interaction with the end user are ignored within this system because there is no method that it should include the users who are familiar with the domain and user interaction must be central part of architectural system. probabilities based learning included to the learning structure which includes that ranking provided according to the system can be more efficient by presenting user to learners system.

We collectively create resources and then provide a compressive introduction to wiktionary in order to take knowledge from wiktionary the structure of wiktionary articles must be evaluated for the inconsistencies among the wiktionary data.hence we will explore how large amount of knowledge can be harvest and how this knowledge is transformed into ontological structure and this structure consist of concepts and relations.

2. Literature Survey

In [1] M.A. Hearst States that in order to avoid need for preen coded knowledge and to provide applicability in wide range of text for this there is need of discovery pattern which can be achieve by using hyponyms relation which satisfies the following things:

In many text genres there occurrence is very frequent. They used to indicate relationship of interest. There is no need of pre-encoded knowledge. In [2] K.T. Frantzi, defines since technical terms are important elements for digital libraries so, here we are presenting a domain independent method with the help of which multi word terms can be automatically extracted. So, it combines C-value and NC-value method in

which C-value method is used for enhancing statistical measure for term extraction with the frequency of occurrence and NC-value cause's extraction of terms from incorporation of information. So, this indicates that C-value and NC-value is domain independent method for extracting multi-word terms using special language English corpora.

J.R. Anderson, finds that in order to develop learning system[3]with the help of [2]Since learning system require metadata and metadata is nothing but the data about data this data is always associated with the learning object manually or it can be partially generated by the system but manual creation of metadata is feasible in small deployment but in case of large deployment it fails because it is not possible to manage it as number of user increases.So,if this is a requirement of learning system then user will provide much more easily a great number of learning object and hence concentrated on metadatasources,document,cotext analysis, document usage, composite document structure. So, there are limited number of pedagogical relations are introduced and lack of ontology relation in the learning object and hence M. Larran˜ aga, concentrated on [5].

In [3] Since learning system requires metadata and metadata is nothing but the data about data this data is always associated with the learning object manually or it can be partially generated by the system this manual creation of data is feasible in small deployment but in case of large deployment it fails because it is not possible to manage as multiple users increases so, If this is requirement of learning management system then user will provide great number of learning object. In[4] So, there was need for automatic generation of generating domain ontology and then use this ontology in computer based education as a domain module and hence, to extract concept maps from text we use a tool called as knowledge puzzle ontology learning tool with the generated domain ontology and then use of three dimensional evaluation advocates:

Structural, semantic, and comparative analysis out of which in structural analysis we provide evaluation in the form of graph, in semantic we provide evaluation based on human expert judgment and in comparative evaluation compares state of art tool and new tool for marking improvements in new techniques. So, although the learning system developed but there was need of reusing learning objects and hence [9].

Hence, A. Zouaq and R. Nkambou, states that[4]there is lack of reusing learning object which causes increase in time complexity and hence he developed [5] there was a need of reusing learning object by retrieving will help in lighten the workloads of construction of new on line courses and hence, this paper presents generation of learning object from electronic textbook by using methods and validation techniques and hence, electronic textbook includes pdf,rtf,and other documents and hence for this used NLP techniques and hence, they include pedagogical relationship which act as special kind of relationship for identifying relation according to requirement. Since [6] the approach that was presented [5] based on ontology but the existing ontology suffers because Of limited ontology learning system, their small size and an unaffordable construction

cost. So, main objective is to introduce wiktionary which is called as online dictionary with the help of which we get encoding information about the words, words relations for the ontology construction.

Hence, it took place in two steps:
1) From wiktionary harvesting of structured knowledge takes place.
2) Ontologies of the knowledge take places that after satisfying with the user or author the final wiktionary construction and make it publicly available and include integration of novel resource with linked data cloud as well as other existing ontology projects.

Since we know that ontology is of good quality but at the same time it contains a large number of concept and hence size is increases because of the ontological data that can be browsed using the web based user interface. but making the ONTOLOGY wiktionary we want more research and hence we are concentrating on the integration of existing ontologies well as linked as data cloud. [10]Bing lieu say that to support the exapansion of web so that to provide richer and richer content and provide user a learning way because previously that was challenging to the traditional method of learning since, in traditional method there was need to find a topic and then survey associated with that topic and hence it was inconvenient for someone to compile all existing knowledge to write a book and there was lack of time to perform all those things .on the other hand when someone try to learn new topic he needs to find definitions of the topic and the subtopic. Hence this paper there is a table of content of subtopics and content pages. Hence this approach called as compiling a book on the web.

3. Proposed Work

The main aim of system is to enable student to learn by themselves in case of lack of other learning system or should provide them guidance about education system in meaningful and isructivist way. So construction of domain module includes
- Textbook preprocessing
- Gathering LDO
- Gathering LOs from document
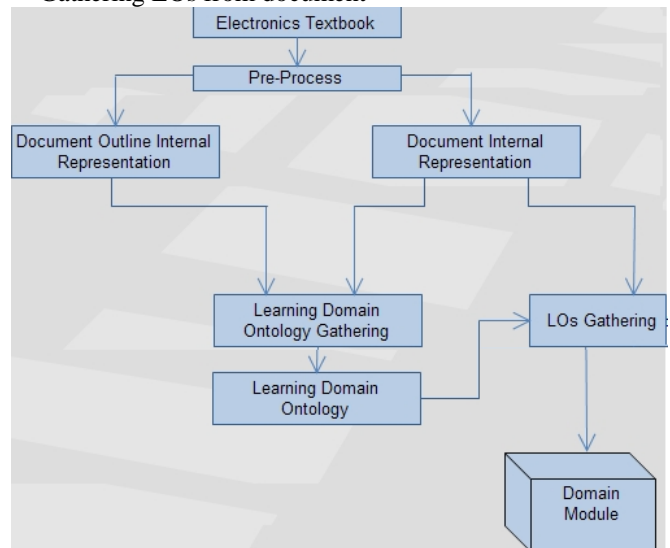
2.1 Textbook preprocessing

For knowledge acquisition process we prepare the document first and then outcome are used to gather knowledge encoded at two levels in the Domain Module.

2.2 LDO gathering

At this stage pedagogical relationship between domain topic which is mastered should be identified and represented into the form of LDO which helps Technology supported learning system to allow student to guide themselves during learning sessions.

2.3 LOs gathering

Examples, definitions, exercises used during learning process are identified. Each step described in detail:

a) *Textbook preprocessing*

In order to run knowledge acquisition process we have to construct electronic document and gathers standardized representation of it but as this documents are available in many different format like pdf, rtf, doc, or odf this preprocessing is important. We use hierarchical structure to organize content of the document ex. it will include hierarchy as documents which includes chapters which further includes sections.

b) *Gathering the LDO*

In this case pedagogical relationship between the domain topic and in stored by LDO so, pedagogical relationship includes structural relations i.e., isA, part of, prerequisite, next in which P isA Q relationship indicate that P is a type of Q.P partOf Q indicate that P is a part of Q, P requisite Q indicate that P must be master to teach Q.

c) *Outline analysis*

It is composed of two main phases:

Basic analysis
For outline internal representation main topic of domain and relationship among these topics are mined from the outline internal representation. So index item is considers as main topic and sub item which describes part of it so structural relationship is described between item and subitem.In this case outline item indicates sequence of learning the domain topics.

Heuristic analysis
It mines new relation based on the previously refined set of heuristics. Heuristic for Structural Relationships:

It allows to identifying kind of relationship between an item of outline and its sub items it works on analysis that only one kind of relation exist between an item and sub item and maximum time it is happened that relation which is to be found is a kind of part of relation.

In this case we observed that:

1) There exist isA relationships between the outline item and all its sub item if group heuristics triggers.
2) Otherwise for individual sub item an individual heuristic that matches will be applied.

Advantages of proposed system:
Existing system works effectively only if input documents are presented appropriately but proposed system works effectively on all kinds of documents. More pedagogical relations are to be identified.

LOs gathering
In this case LO-definations, examples, exercises and so on to be identified during the learning process which includes processing of electronic textbook after processing it apply learning domain ontology which generates learning object which again generates DR grammar which is used to find text fragments that might contain appropriate resources [11]. So, we have to convert from DR to LOs and hence, once the fragments are generated of resources again objects are build which are stored in zip file and zip file contain all those storage in XML format.

4. Conclusion

This paper utilizes that how domain module is generated from electronic textbook which is provided in different forms of document so that time complexity can be reduce as there is no requirement of domain specific knowledge and hence it is independent. This paper improves pedagogical relationship as there was lack of basic analysis and outline analysis in previous paper.

References

[1] M.A. Hearst, "Automatic Acquisition of Hyponyms from Large Text Corpora," Proc. 14th Conf Computational Linguistics (COLING '92), pp. 539-545, 1992.

[2] K.T. Frantzi, S. Ananiadou, and J. Tsujii, "The C-Value/NC-Value Method of Automatic Recognition for Multi-Word Terms," Proc. Second European Conf. Research and Advanced Technology for Digital Libraries (ECDL '98), pp. 585-604, 1998.

[3] M. Meire, X. Ochoa, and E. Duval, "SAmgI: Automatic Metadata Generation v2.0," Proc. World Conf. Educational Multimedia, Hypermedia, and Telecomm. (ED-MEDIA '07), pp. 1195-1204, June2007.

[4] A. Zouaq and R. Nkambou, "Evaluating the Generation of Domain Ontologies in the Knowledge Puzzle Project," IEEE Trans. Knowledge and Data Eng., vol. 21, no. 11, pp. 1559-1572, Nov. 2009.

[5] M. Larran˜ aga, A. Conde, I. Calvo, A. Arruarte, and J.A. Elorriaga,"Evaluating the Automatic Extraction of Learning Objects from Electronic Textbooks Using Erauzont," Proc. 11th Int'l Conf. Intelligent Tutoring Systems (ITS '12), pp. 655-656, 2012

[6] Semi-Automatic Ontology Development: Processes and Resources, M.T. Pazienza and A. Stellato, eds., IGI Global, 2012.

[7] Ontology Learning from Text: Methods, Applications, and Evaluation, P. Buitelaar, P. Cimiano, and B. Magnini, eds., IOS Press, 2005.

[8] T. Leidig, "L3-Towards an Open Learning Environment," ACM J. Educational Resources in Computing, vol. 1, no. 1, pp. 5-11, 2001.

728

[9] K. Verbert, "An Architecture and Framework for Flexible Reuse of Learning Object Components," PhD dissertation, Faculteit Ingenieurswetenschappen, Katholieke Univ. Leuven, Feb. 2008.

[10] B. Liu, C.W. Chin, and H.T. Ng, "Mining Topic-Specific Concepts and Definitions on the Web," Proc. 12th Int'l Conf. World Wide Web (WWW), pp. 251-260, 2003.

[11] M. Larran˜ aga, J.A. Elorriaga, and A. Arruarte,"A Heuristic NLP Based Approach for Getting Didactic Resources from Electronic Documents," Proc. European Conf. Technology Enhanced Learning (EC-TEL '08), pp. 197-202, 2008.