

Automatic Identification of Learning Object Type

Sonal Jain
Associate Professor
GLS Institute of Computer
Technology-MCA
Ahmedabad-58

Jyoti Pareek
Department of Computer
Science, Gujarat University,
Ahmedabad-06, India

ABSTRACT

With energetic development of the Internet, online learning object repositories and web based learning systems have become more and more realistic and popular in the past ten years. One of the essential characteristics of learning object is, it should be discoverable depending on its characteristics. This usually entails tagging them with appropriate descriptive metadata. Learning is a cognitive activity that differs from student to student. Achieving personalization and adaptability is not possible till the learning objects are properly attributed. Again metadata instantiation is a difficult task. The Dublin Core metadata element set and IEEE Learning object metadata are two popular metadata standards that facilitate tagging of learning material. This paper showcases automatic identification of Learning Object Type. Experimental evaluation and results depict satisfaction level of users.

General Terms

Information Retrieval

Keywords

Personalization, Learning Object Metadata, Learning Object Type

1. INTRODUCTION

The Learning Object Type attribute is included in IEEE LOM (5.2) under educational category and its explanation as per IEEE 1484.12.1-2002 Final Draft Standard is “Specific kind of learning object” with most dominant kind shall be first” [7]. The type field of Dublin Core metadata set also suggests the same [3]. A Learning resource can be attributed as Explanation, Application, Experiment, Exercise and/or Case study. Every learner has different set of requirements and his/her learning requirements may vary depending on his subsequent learning stages. Sometimes s/he may be looking for explanation of a concept and at times s/he may require exercises on the same concept. For example, if a learner seeking case study on Six Sigma, mentions “Six Sigma” as searching text, s/he will find list of all the documents containing Six Sigma. Instead, if searching can be précised by specifying exercises on Six Sigma, results would lead to retrieval of more relevant documents. Search engine can look for Learning resource type stored as part of metadata and can result into data retrieval depending on the need of user. Recommender systems can present the learning document in a certain sequence using Learning Resource Type metadata attribute. For example, at a first stage, a material explaining the topic/concept can be presented to a learner, in second stage applications of the topic/concept can be presented, at a third stage, case study/experiment/example of topic/concept can be presented and at the fourth stage exercise on the same concept to help learner to examine his knowledge gained on that particular concept can be presented to learner. This

sequence can beyond doubt enhance the learning experience of a learner. This can be achieved by storing learning resource type for each learning document uploaded in the system.

However, once the metadata to be stored is identified, the next question arises is who will enter this metadata. Manual annotation is a time consuming and expensive process [8] [4] [5]. Correct instantiation of Learning Object metadata requires combined educational and technical skills. It is also liable to human errors [6]. Many people feel [11] [10] [12] that unless the process of annotating learning objects is automated, it is difficult to create a critical mass of reusable learning objects. Human tagging works well only in situations where the number of participants greatly exceeds the number of resources to be tagged and where there is no requirement for controlled vocabularies or standardized metadata formats [9]. Realizing the need we have worked on the automatic semantic annotation of textual learning materials. A number of methods like parsing and analysis of documents have been used for automatic extraction of the above metadata attributes. The automatic extraction of some of the metadata also makes use of the domain ontology [13] [1]. This paper showcases automatic identification of Learning Resource Type. Section II contains discussion of our approach. User interface, evaluation methodology and results and conclusion and future work are discussed in section III, IV and V respectively.

2. OUR APPROACH

The algorithm developed by us identifies Learning Resource type on the basis of pattern-base created as part of this research work. Fig. 1 shows process flow for Learning Resource type identification.

2.1 Pattern Base

Our observation of learning material states that to define a concept, to apply a concept, to discuss experiments or case study or while designing exercises, certain grammatical patterns are followed. After analyzing large number of documents, certain common patterns were listed down and sentences containing those common terms were considered as candidate for learning resource type extraction. These patterns can be exploited to identify the type of learning document. To identify patterns, set of learning materials from different subject domains like operating systems, database management systems, statistics, total quality management and data structures were collected from various learning object repositories like MERLOT, websites of various professors of

various universities, Wikipedia etc. (Table II) and these documents were thoroughly analyzed and patterns were listed down as shown in Table I. Certain patterns picked by us were also listed as a list of useful verbs used in teaching and learning by Dalton.J. & Smith to apply blooms taxonomy [2].

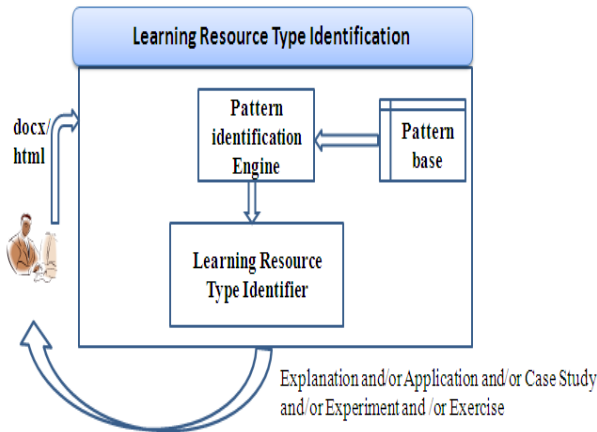


Figure 1: Process Flow for Learning Resource Type Identification

2.2 Pattern Identification Engine

This phase picks up sentences from a document one by one and for each sentence, presence of pattern pre-stored in pattern-base is checked. Separate counters keep track of number of patterns identified under each category. Counter value under corresponding category is incremented if pattern listed under that category is found in the sentence.

Table I: Trigger patterns identified for identification of Learning Resource type

Learning Resource type: Explanation
are, called as, can be classified as, can be explained as, can be viewed as ,can occur ,consists of, contains, deal with, deals with, defined as, depicted as, described as, explained as, holds, interpreted as, is, is a ,is an, based on, called a, called an, called as, called the, defined as, defined by, described as, intended to, made up of, referred as, referred by, termed as, viewed as, known as, manages, means, occurs, recognized as, referenced by, referred by, refers to, represent an ,represents the, says, stands for, termed as, used as, used for, used to
Learning Resource Type: Application
application of, used for, applied for, apply, demonstration, illustration
Learning Resource Type: Experiment
achieve, results, objective, experiment ,findings, method, steps, experiment, evaluation
Learning Resource Type: Exercise

Analyze, answer-in-brief, choose, compare, compare and contrast, contrast, convert, criticize, define, describe, differentiate, discuss, distinguish, enumerate, evaluate, evaluate, explain, fill in, find, give reason, highlight, highlight the differences between, how, how can, how many, identify, illustrate, improve, interpret, justify, label, list, locate, make a list of, match the following, name, outline, persuade, point, predict, prepare, prove, reconcile, relate, restate, review, select, solve, state, suggest, summarize, tabulate, tell, trace, translate, underline, what, when, where, which, write, who

Learning Resource Type: Case Study

Case study

2.3 Additional Rule for Exercise Identification

In case of Exercise, additional constraint is placed. Counter value corresponding to Exercise will be incremented only if pattern listed under that category is found in the beginning of the sentence.

2.4 Additional Rule for Case Study Identification

For case-studies, it was observed that very less number of patterns gets repeated. But it was apparent that case study consists of collection of past tense sentences. Thus additional rule was formed for case study. A case may arise where any of the patterns listed in Table I is not found in any sentence, however 80% or more sentences in learning material are in past tense, it implies that it can be considered as case study.

2.5 Additional Rule for Application Identification

A learning document may belong to multiple topics/subtopics which may in turn belong to multiple subjects or disciplines [13]. It was observed that learning document contains application of the concepts which come under the topics which is not the most significant topic discussed in the learning material. A learning resource was found to contain application of certain concepts if it belongs to multiple topics. Thus learning resource can be considered as an application if it is associated with multiple topics/subtopics.

2.6 Learning Resource Type Identifier

This phase examines counter values used for each Learning Resource type, which contains number of patterns identified under each resource type category. Categories are stored in Learning Resource Type metadata attribute and stored in descending order of counter values of each category. It finally lists down learning resource types of which corresponding counter values are greater than zero

3. SAMPLE INPUT AND OUTPUT

The learning material can be provided as input as shown in figure. 2 and Sample output would be as shown in figure. 3.



Figure 2: User Interface

Semantic Annotation of Learning Objects (SALO)	
Document Details:	
Document Name	cpuScheduling.htm
Document Title	CPU/Process Scheduling
Granularity Level	2
Learning Resource Type	Explanation Application

Figure 3: Sample Output

4. EVALUATION AND RESULTS

In this study, we have randomly collected sample of 200 learning material belonging to different subjects like Operating systems, Database management systems, Data and File structures, Statistics and Total quality management. The documents were processed by our tool and learning resource type for each document was extracted by tool. For evaluation, subject experts were given the extracted outcome by our tool. Subject expert's agreement level was sought for each extracted learning resource type. Figure 4 shows sample of subject author's agreement/disagreement with respect to system output.

When extracted concepts were tagged by authors with their agreement/disagreement, it was found that they agreed in 77.78% cases of Explanation, 60% in case of Application, 89.23% in case of Case Study, 55.32% in case of Experiment and 78.57% in case of Exercise as shown in figure 5.

Document Title	Learning Resource Type	Agree/Disagree
achievemutualexclu.htm	Explanation	Yes
	Application	Yes
sixsigma.htm	Explanation	Yes
	Application	Yes
standarddeviation.htm	Explanation	Yes
	Application	Yes
centrallimtheorem.thm	Explanation	Yes
	Application	Yes
virtualmemory.htm	Explanation	Yes
	Application	Yes
uni_paper_2006.docx	Explanation	No
	Exercise	Yes
paging.docx	Explanation	Yes
	Application	Yes
TQM.htm	Case Study	Yes
process.htm	Explanation	No
	Experiment	No
cpuScheduling.htm	Explanation	Yes
	Application	Yes

Figure 4: Sample of Subject Expert's Agreement with System outcome

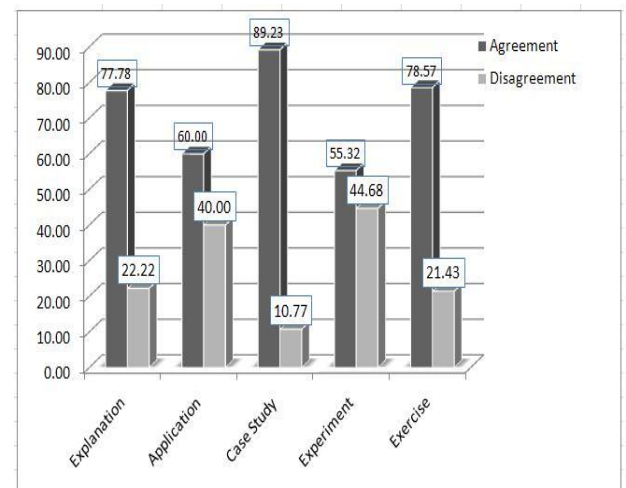


Figure 5: Evaluation results summary

5. CONCLUSION AND FUTURE WORK

This paper includes approach for automatic identification of Learning Resource Type. Identified Learning Resource type includes Explanation, Application, Exercise, Experiment and Case Study as part of this research work. Experimental Evaluation and results show that proposed algorithm is effective in identifying learning resource type.

It may be noted that a learning material may be explanation of one concept as well as application of other concept. A learning material may be Experiment as well as can be considered as a Case Study. We suggest that along with the learning object type, further details should be stored. As discussed earlier, learning material may contain explanation of one concept and at the same time application of another concept. Now in this case a learner seeking explanation of "Standard Deviation" if searches the documents by mentioning "Standard Deviation" as a searching word, the documents containing explanation of Six Sigma would also be displayed. If only learning object type is stored as part of

metadata, learning material containing application will be listed by search engines and recommended by recommender systems only when learner specifically fires “application” in search criteria. If along with the Learning object type, explanation of which concept and application of which concept is bifurcated, recommender systems can intelligently present application of concepts when a learner either has grasped the explanation of the concept or is trying to understand the concept. This recommendation of application at the time of learning the concept will make the concept clearer in learner’s mind as he will be able to relate theory as well as application at the same time. Thus our future work includes identification of concepts with respect to each learning resource type tags associated with the document.

6. REFERENCES

- [1] B. Chandrasekaran, J. R. Josephson, V. R. Benjamins, (1999), What are ontologies, and why do we need them?, *IEEE Intelligent Systems*, 14(1), pp. 20–26
- [2] Dalton, J. & Smith, D. (1986), *Extending Children’s Special Abilities – Strategies for primary classrooms*, pp. 36-37
- [3] DublinCore, <http://dublincore.org/>
- [4] Duval, E., Hodgins, W. (2004), *Metadata matters. Proceedings of International Conference on Metadata and Dublin Core Specification, DC-2004, Shanghai, China.*
- [5] Friesen, N. (2004), *International LOM survey: Report. ISO/IEC JTC1/SC36 subcommittee*, http://mdlet.jtc1sc36.org/doc/SC36_WG4_N0109.pdf.
- [6] Hatala, M., Forth (2003), S. *System for Computer-aided Metadata Creation*, In *Proceedings of 12th International Conference of The World Wide Web Consortium (WWW2003)*, Budapest, May 20-24.
- [7] IEEE LTSC, *IEEE Learning Technology Standards Committee*, <http://www.ieeelts.org>
- [8] Leacock, T., Farhangi, H., Mansell, A., Belfer, K (2001), *Infinite Possibilities, Finite resources: The TechBC Course development Process*, *Proceedings of the 4th Conf. on Computers and Advanced Technology in Education (CATE 2001)*, June 27-29, Banff, Canada, pp.245-250.
- [9] Marko a. Rodriguez, Johan Bollen, Herbert van de sompel (2009), *Automatic Metadata Generation Using Associative Networks*, *ACM Transactions on Information Systems*, Vol. 27, No. 2, Article 7.
- [10] M. Meire, X. Ochoa, and E. Duval (2007), *SAMgI: Automatic Metadata Generation v2.0*, *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, pp. 1195-1204.
- [11] Ochoa, X., Cardinaels, K., Meire, M., & Duval, E. (2005), *Frameworks for the Automatic Indexation of Learning Management Systems Content into Learning Object Repositories*. *World Conference on Educational Multimedia, Hypermedia & Telecommunications, EDMEDIA 2005*, Montreal, Canada, pp. 1407-1414.
- [12] Paramjeet Singh Saini, Marco Ronchetti, Diego Sona (2006), *Automatic Generation of Metadata for Learning Objects*, *Sixth IEEE International Conference on Advanced Learning Technologies*, pp.275-279.
- [13] Sonal Jain, Dr.Jyoti Pareek, *Automatic Topic(s) Identification from Learning material: An Ontological Approach*, *IEEE Explore*, Print ISBN 978-0-7695-3982-9, INSPEC Accession Number: 11242737, Digital Object Identifier: 10.1109/ ICCEA.2010.221

Table II: Websites used for collection of learning documents and identifying patterns

- <http://www.personal.kent.edu/~rmuhamma/OpSystems/os.html>
- <http://www.wikipedia.org/>
- <http://stattrek.com/>
- <http://www.merlot.org/merlot/index.htm>
- <http://cnx.org/>
- <http://www.fredosaurus.com/notes-db/index.html>
- <http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/>
- <http://www.learning-objects.net/>
- <http://pages.cs.wisc.edu/~bart/537/lecturenotes/titlepage.html>
- <http://lass.cs.umass.edu/~shenoy/courses/fall00/lectures.html>
- <http://people.csail.mit.edu/rinard/osnotes/>
- <http://redbook.cs.berkeley.edu/redbook3/lecs.html>
- <http://www.cs.jcu.edu.au/Subjects/cp1500/1998/foils/introDB.html>
- <http://www.cs.umb.edu/~poneil/dbppp/notes/>
- <http://content.hccfl.edu/facultyinfo/tachenbach/statsnotes.html>
- <http://people.richland.edu/james/lecture/m170/>