An Effective Method for Semantic Richness Assessment of Electronic Content by Using WordNet Ontology

Ahmad A. Kardan¹, Shabnam Teimoori²

¹²Advanced Technologies in E-Learning Laboratory
Amirkabir University of Technology, Iran
E-mail: teymuri_sh@yahoo.com

Abstract
Content is one of the important elements of any electronic environments. Therefore, Web content can be the best medium for communication and collaboration. Increasing the importance of content can be investigated from two aspects: firstly from the search engines point of view and secondly, from the users’ perspective. During the search, engines use keywords to find quality content and this is the main criterion for their ratings. Therefore, adding keywords is an effective method which is used by internet experts. In e-learning, the ability to measure the effectiveness of learning is important since it can be used to support educational objectives. Despite the fact that many systems are to evaluate the content, they have not yet used a developed methodology for assessing the e-content. In this study, we examine the tags which are attached by the user in order to evaluate the content. The more complete is the content in semantic richness, the more accurate users and search engines can identify the key concepts. In this paper, an evaluation architecture and algorithm have been suggested for the aforementioned purpose. Moreover, its implementation and its test are carried out in an e-learning system. In this study, Word sense disambiguation strategy based on the lexical WordNet database is considered to determine the accuracy of the tags assigned by the user in comparison with an expert. At the end, the outcome of the evaluation is compared with an expert’s judgment. The results indicate good accuracy and precision of the proposed solution.

Keywords: Electronic content, assessment system, semantic richness, word sense disambiguation

1. Introduction
Content assessment make it possible to restructure the data and it is a qualitative and quantitative study. As the electronic content is in multiple formats, its evaluation encounters many challenges (Neuendorf, K.A, 2002).

In e-learning environment, the importance of content evaluation is related to the underlying fact that it has a direct effect on the learner's learning ability. In educational environments and e-learning in particular, evaluation of the teaching ability of content as a means can be utilized to advance educational goals (Kim and Kuljis, 2010). This means prior to the end of a training period, the course teaching effect can be assessed and if necessary, the required changes should be applied. There are different criteria for evaluating the content; among them are: The validity, reliability, accuracy, availability scalability, comprehensiveness and teaching ability (Palmquist, 2002).

Evaluation of content quality and received information by the user is not simply possible for him. Quality assessment of information is also important to create knowledge from information in training (Steinberg, Brehm, 2010).
Among the evaluation criteria, assessing content integration and its teaching ability, despite its critical importance, are less examined.

Our goal is to extract the hidden knowledge of user’s interactions with the content and use it to evaluate the content. For this purpose, architecture and an evaluation algorithm were presented; moreover, a system was implemented where electronic content was presented to user and he/she was asked to place tags on the realized key concepts of the content. Evidently, the more teaching ability the content has, the easier is to identify the concepts by users.

2. Content Assessment
Evaluation of electronic content is important in e-Learning because there is no direct relationship between teacher and learner; thus it is essential that learner is helped to find appropriate and essential information.

Content assessment considered of the various aspects include: Quality content, content aesthetic, and organizing content. Quality content includes: the relevancy, accuracy, being up to date, completeness, semantic richness and etc. Content aesthetic is from the perspective of visual and understandable format. Organizing content is accompanied with the goal of page structure as well (Steinberg and Brehm, 2010). In the area of quality content, we assess comprehensibility (semantic richness) and completeness of the content which it leads the content teaching ability to be increased.

2.1 Assessing Semantic Richness of Content
Semantic richness refers to meaningful information which is related to content. This evaluation item can be done by means of presenting knowledge tools such as ontology and WordNet. Assessment is divided broadly into two categories, formative and summative. Formative evaluation conducted during the learning phase with goal of promoting learning. Summative evaluation conducted after the learning phase, with goal of providing suggestions for improving curriculum (Adline and Mahalakshmi, 2011).

The evaluation in this research is formative assessment during the learning period with the aim of improving learning.

2.2 Word Sense Disambiguation and WordNet Dictionary
Human language is ambiguous, so that some of the words in the content are interpreted in different ways. Computational identification of the meaning of words in text is called word sense disambiguation (WSD).

Word sense disambiguation is a process for determining the meaning of words in a particular text. A text T can be regarded as a series of words and consider WSD as a function for allocating concepts to some or all of the words in T; thus a mapping of words to concepts specifies that

\[ A(i) \rightarrow \text{SenseD}(wi) \]

In equation 1, ‘Sense D (wi)’ is decrypted meaning of the words in D dictionary for ‘wi’ word and \( A(i) \) is a subset of ‘wi’ meanings that are appropriate for \( T \) text. "A" map can be assigned to more than one meaning to each word. While the best meaning is chosen which means: \( |A(i)| = 1 \).

2.3 Selecting the Relevant Sense
Sense of the word is an accepted meaning of the word. WSD aims to clarify the meaning of the words in a text by the processing method. A method to determine the most suitable meaning of words, concepts, is to extract a graph structure from the words meaning which is associated with
the theory of lexical chains. A lexical chain is the sequence of words \((w_1, w_2, \ldots, w_n)\) in a text which are semantically related so that \(w_1\) is connected to \(w_{i+1}\) by a lexical semantic relation such as (is-a, has-part, ...); similar to the following chain meaning: eat \(\rightarrow\) dish \(\rightarrow\) vegetable.

In general, we can consider two general methods for WSD. Controlled WSD methods that use machine learning techniques in the code samples labelled to determine the meaning and uncontrolled WSD methods are based on a set of unlabelled texts. In addition, there are knowledge-based techniques (based on dictionary and rich knowledge). In this study, we used the method based on dictionary. Glossary WordNet is our choice. WordNet dictionary is a language processing dictionary which is based on the psycholinguistic and has been created at Princeton University. WordNet dictionary codes the concepts as a set of synonymous words which the same concepts are a unique word (Semeraro et al., 2007).

2.4 Synonym Word Similarity Measures
Since the early 1990s, the introduction of WordNet, criteria have been developed to determine the semantic similarity of concepts in order to extract semantic network of word relationships. Semantic similarity measures are defined in equation 2:

\[
\text{SCORE} : \text{SENSE}_D \times \text{SENSE}_D \rightarrow [0,1]
\]

Senses D is a complete set of concepts in the reference dictionary. The number of points in equation 2 is a value between 0 and 1. A similarity measure based on the distance between \(S_w\) and \(S_{w'}\) has been introduced by Leacock and Chodorow. In relation (3), \(d(S_w, S_{w'})\) is the number of nodes in the shortest path from the hierarchical structure of IS-A classification hierarchy WordNet and D is the maximum depth. This is our criteria in determining the similarity of user’s tags to those determined by the expert (Lee et al., 2009).

\[
score_{Lch}(S_w, S_{w'}) = - \log \frac{d(S_w, S_{w'})}{2D}.
\]

3. Proposed Architecture of Content Assessment
Hitherto, no practical method is implicitly presented to evaluate e-learning content which is collected from the interaction between a user and content. For instance in (Adline and G.S, 2011), a general architecture is considered for content evaluation which despite including various aspects of evaluation, no method is expressed for implementation and practical use. In (Seetha, 2012), a system is presented for content evaluation which scores the content quality via collecting information and learners’ opinion in universities and experts’ sites. The advantage of this method is to collect information from various sources and to categorize them. However with respect to explicit evaluation, learners may not always cooperate essentially.

As it is evident in figure 2, system scenario is described in four phases: content creation Layer and, the user interface layer, content assessment layer, and recorded results layers. The presented algorithm is illustrated in figure 1 section B.

3.1 Content Creation Layer
In this layer, content writer will adjust electronic content and insert it to the system. (Figure 1 section A, at top of the figure). Educational content for quick search and efficient management are generated and logged in the form of learning objects (LO). The LO are independent and reusable learning resources. According to the IEEE Learning Technology Standards Committee (IEEE/LTSC), a LO can be defined as "any entity, digital or non-digital, that may be used for learning, education or training" (Wu and Doulai, 2009).
3.2 User Interface Layer

As shown in figure 1 this layer is part of the definition of system users and their access levels. The educational content that was created by a teacher in the previous layer is being available for learners and experts of the system.

Learners study the educational content and put tags on them. These tags are the main concepts of the lesson that learners have learned them. Similarly, the lessons that were tagged by the learners will be available for experts.
The expert’s task is to determine the main concepts of the lesson which are identified as keywords. Each module consists of a set of the main concepts. Expert provides the knowledge needed for the evaluation by determining the key words.

3.3 Content Creation Layer
In this process, the user is asked to place tags on the concepts he/she has learned from the lesson as the keywords of each lesson on their tags. On the other hand, the expert determines the related concepts of the content. To determine the meaning of a word in the WSD, word text is required. Consequently, we need the text of tag to define the meaning of that ambiguous tag. To find the ambiguous tag content, its adjacent tags can be used as the text. This idea is called as tag sense disambiguation and its logic is that frequent co-occurrence of two tags represent the semantic relationship between them. In this project, the key words determined by an expert specify the adjacent tags.

3.4 Recording Result Layer
In this part, the extracted knowledge from the evaluation layer is inserted to the evaluation database. The teacher will compare the results of the evaluation and will make the required course modifications and the lesson will be given to the new users for repeated evaluation.

4. Implementing the Architecture
System has been implemented with these modules and tools as expressed below: Tomcat application server, Jsp pages, J2ee framework, WordNet java tool.

5 Testing the Proposed Method
To test the proposed evaluation method in e-learning, a web site was designed and implemented. In this system, the learners were logged and courses under evaluation were given to them. Elective courses were in IELTS vocabulary instruction field.

5.1 Participants in the System
In this system, two experts in the field of education (teacher training IELTS test) cooperated. In the first phase, the primary tests were performed to determine the level of learners. Subsequently, the remaining members are divided into three categories of: elementary, intermediate and advanced (groups A and Band C) according to the results of primary test. Next, each of these groups was studied the same educational content which includes 5 courses. For the final evaluation, 21 active users were used.

In table 1, the heading row indicates courses title and heading column presents training groups and experts. In table cells, the average number of correct tags for each lesson is provided (row 2 to row 4). From row 5, both experts’ estimates are presented for correct tags of each group for 5 courses.

<table>
<thead>
<tr>
<th></th>
<th>Course 1</th>
<th>Course 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group A</strong></td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td><strong>Group B</strong></td>
<td>1.7</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Group C</strong></td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>Group A – Expert 1</strong></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><strong>Group A – Expert 2</strong></td>
<td>2.5</td>
<td>4</td>
</tr>
<tr>
<td><strong>Group B – Expert 1</strong></td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Group B – Expert 2</strong></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>Group C – Expert 1</strong></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Group C – Expert 2</strong></td>
<td>1.5</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 2. Comparison of error percentage for users and expert for each lesson

<table>
<thead>
<tr>
<th></th>
<th>Course 1</th>
<th>Course 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Percentage for Users and Expert 1</td>
<td>33.3</td>
<td>6.66</td>
</tr>
<tr>
<td>Error Percentage for Users and Expert 2</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Average</td>
<td>11.66</td>
<td>13.33</td>
</tr>
<tr>
<td>Total Average</td>
<td></td>
<td>12.49</td>
</tr>
</tbody>
</table>

Table 2 compares the error percentage in terms of users and the expert’s opinion. In the first row, the error percentage rate was estimated with the first expert’s opinion; moreover in the second row, the error percentage is compared with the second expert’s estimate. In the third row, the average error percentage with two experts’ opinions is presented. We use equation 4 to calculate the error percentage:

\[
\text{ErrorPercentage} = \frac{\text{Analysis} - \text{ExpertOpinion}}{\text{ExpertOpinion}} \times 100
\]

On average, there is %12.49 of error for all the courses assessment. In figure 2, the number of correct tags for each learner and for 5 evaluated lessons is displayed by the graphs. As is evident from the graph, the average rate of correct tags from Lesson 1 to Lesson 5 is increased. This is because we have increased the comprehensiveness of content from Lesson 1 to Lesson 5.

The dark blue colour presents lesson 1, red colour indicates second lesson, green colour is assigned for third lesson, the purple colour represents fourth lesson, the blue light stands for the fifth lesson, the number of correct tags is on the right vertical axis and the horizontal axis shows the users. The red oval on the diagram typically represents the increase in the number of correct tags for the 14th student which his/her number of tags has been increased from lesson 1 to lesson 5.

Figures 2. Chart of the correct tags on each lesson

In figure 3, the blue shows group A, the red colour represents expert 1, the green colour is assigned for expert 2, the vertical axis shows the number of correct tags estimated by experts and evaluated by users. The horizontal axis represents the lessons which have been provided. As the diagram shows, there is a very little difference between the learners and the expert. Moreover by calculating the error percentage for 5 courses, an average for confidence factor is obtained. The error percentage is calculated as: 15.58 by equation 4.
6. Conclusion

In this paper, a method of assessing electronic content was presented through recording user’s interactions with the content. In this study, a tagging environment was created which its aim is to recognize the rate of comprehensive and teaching ability of content. Finally, a method was presented which requires no continuous need for an expert to assess content. This method evaluates content with respect to users’ opinions.

Due to the fact that learner tags the content based on him/her understanding and machine need methods and algorithms to recognize its meaning, natural language processing approach is applied to determine the accuracy of the user in the context. This is called the method of word sense disambiguation.

This method relies on lexical linguistic stored knowledge in the WordNet ontology. The Leacock-Chodorow algorithm is a method of knowledge-based WSD. This algorithm defines a measure of similarity between the concepts based on the ontology hierarchy of wordNet.

7. References

7.1 Books:

7.2 Journal Articles:

7.3 Conference Proceedings:
Inhwa Kim and Jasna Kuljis.( June 21-24, 2010): Applying Content Analysis to Web based Content, 32nd Int. Conf. on Information Technology Interfaces, Croatia IEEE.

7.4 Internet Sources: