ABSTRACT

As text sources are getting broader, measuring text similarity is becoming more compelling. Automatic text classification, search engines and auto answering systems are samples of applications that rely on text similarity. Learning management systems (LMS) are becoming more important since electronic media is getting more publicly available. As LMS continuously needs content enrichment and the web is getting richer, automatic collection of learning materials becomes an innovative idea. Intelligent agents can be used with a similarity measurement method to implement the automatic collection process. This paper presents a new method for measuring text similarity using the well-known WordNet Ontology. The proposed method assumes that a text is similar to another if it represents a more specific semantic. This is more suitable for LMS content enrichment as learning content can usually be expanded by a more specific one. This paper shows how the hierarchy of WordNet can be taken advantage of to determine the importance of a word. It is also shown how similarity method within an e-learning system is exploited to achieve two goals. The first one is the enrichment of the e-learning content, and the second is the detection of semantically similar questions in e-learning questions banks.

Keywords: Intelligent Agent, Learning Management Systems, Semantic Similarity, Text Similarity, WordNet

INTRODUCTION

The web is getting broader by the day and richer contents are getting more available. However, browsing the entire web to collect all useful content is an intractable mission for human beings. So, automatic text similarity bots can be used to search the web for relevant documents.

Much research works have been carried out in this field in the last two decades; some of them employ statisti-
cal methods which are based on pre-
classified terms extracted from a corpus. 
Others depend on semantics and use the 
natural language processing techniques.

We present in this paper, a new se-
monic method for text similarity mea-
surement based on WordNet Ontology
and suitable for learning management
systems. Section II shows a brief sum-
mary of some previous related works.
Sections III and IV describe the proposed
method and its algorithm. Section V
shows how the method is applied to en-
rich LMS content. Section VI describes
how to use the similarity method to detect
similar questions in questions banks.
The results are reviewed in section VII
and the paper is finally concluded with
section VIII.

LITERATURE REVIEW

Many methods have been presented
to measure text similarity. Traditional
methods are based on text lexical anal-
sis and adopted by many information
retrieval systems to find similar texts
based on a text query. Some new research
works are based on corpora-extracted
statistics, and are considered to be sta-
tistically oriented (Mihalcea, Corley, &
Strapparava, 2006; Corley & Mihalcea,
2005; Islam & Inkpen, 2008; Amala Bai
& Manimegalai, 2013). Many other stud-
ies have focused on the concepts of texts,
where some conceptual representations
like ontologies is used to determine the
overwhelming concepts of a text (Pandya
&Bhattacharyya, 2005; Wang & Taylor,
2007). Some other works are based on
machine learning techniques, where an
agent is used to learn how to test text
similarity (Bilenko & Mooney, 2003;
Lee, Pincombe, & Welsh, 2005). Some
hybrid systems are also proposed such as
the one in Mohle and Mihalcea (2009).

Using Ontologies in e-learning sys-
tems were presented in many researches,
as in Henze, Dolog, and Nejdl (2004)
where the authors proposed a method
to personalize e-learning contents using
Ontologies and semantic web resources.
They investigate a logic-based approach
to educational hypermedia using TRIPLE,
which is a rule and query language
for the semantic web.

Many other researchers used Word-
Net in e-learning, Carbonaro (2010)
proposed a research that aims to build a
summarization system to support tutors
in managing student communication
and interaction within an educational
environment. They show that Concept-
based approaches to represent dynamic
and unstructured information can be
useful to address issues such as trying to
determine the key concepts and to sum-
marize the information exchanged within
a personalized environment. It seems a
promising technology for implementing
distinct learning environments; enabling
the organization to deliver learning materials around small pieces
of semantically enriched resources.

The study in Hung and Yee (2005)
shows a semantic-based automated ques-
tion answering system that can act like a
virtual tutor to answer student questions
online. This system, not only relieves
the tutor from the burden of answering
many questions, but also allows students
to get answers promptly without waiting
for the tutor’s response.

Another research example of using
Ontology in e-learning is Deline, Lin,
Wen, and Gašević (2009). This research
proposed an ontology-driven software
development methodology which is ap-
propriate for intelligent ontology-driven
Related Content

A System for Automatic Evaluation of C Programs: Features and Interfaces
[www.irma-international.org/article/system-automatic-evaluation-programs/2992/](www.irma-international.org/article/system-automatic-evaluation-programs/2992/)

Creating Electronic Evidence Rooms for NCATE Accreditation: Using College LiveText Solutions
[www.irma-international.org/chapter/creating-electronic-evidence-rooms-ncate/44360/](www.irma-international.org/chapter/creating-electronic-evidence-rooms-ncate/44360/)

Perceptions of Preparation of Online Alternative Licensure Teacher Candidates
[www.irma-international.org/chapter/perceptions-preparation-online-alternative-licensure/67970/](www.irma-international.org/chapter/perceptions-preparation-online-alternative-licensure/67970/)

Blogs and Wikis
[www.irma-international.org/chapter/blogs-wikis/30418/](www.irma-international.org/chapter/blogs-wikis/30418/)

Course Management Systems: Hope or Hype?
[www.irma-international.org/article/course-management-systems/2980/](www.irma-international.org/article/course-management-systems/2980/)