Building Domain Ontologies for Hyperlinked Multimedia Pedagogical Platforms

Zenun Kastrati, Ali Shariq Imran, and Sule Yildirim Yayilgan

Department of Computer Science and Media Technology Gjøvik University College, Norway
{zenun.kastrati,ali.imran,sule.yayilgan}@hig.no

Abstract. This paper examines building of the course ontology for describing and organizing hyperlinked pedagogical content. The ontology is used to structure and classify multimedia learning objects (MLO) in hyperlinked pedagogical platform called HIP, and to assist students to search for lectures and other teaching materials in a reasonable time and more efficiently. In addition, this paper proposes a new approach to improve the classification performance by enhancing the information representation model using concepts from the pedagogical course domain ontology. The model will automatically estimate weight of concepts within the ontology, and it will combine the weight with concepts’ importance which is calculated using Term Frequency Inverse Document Frequency – tf×idf algorithm. This paper is a work in progress. We are in process of creating and implementing the course ontology and an experiment will be conducted to evaluate the classification performance in terms of efficiency and effectiveness for the approach proposed in this paper.

Keywords: Ontology, concept vector space, Markov chain model, HIP.

1 Introduction

Recently, the rapid developments in technology and the increasing usage of computer and other electronic devices have made other forms of education possible, which are different from the traditional ones. The distance learning is one amongst them where students can access the information without being present in class. To improve distance learning process, new means of communication and studying were introduced and new frameworks were created [1]. One such eLearning platform is Hyper Interactive Presenter – HIP [2].

HIP is a technology-rich pedagogical platform that uses combination of media elements to deliver the learning objects. The elements comprise of electronic documents such as wiki pages and PDF documents, presentations, lecture videos, an intelligent pedagogical chat bot. In addition, it also provides navigational links, tagged keywords, and frequently asked questions (FAQ).

A huge amount of information in HIP platform comes from different media modalities that need to be organized, structured and hyperlinked. Structuring and organizing such a wealth of information is labor intensive, prone to errors and a cumbersome task. Structuring information is required due to the need for classifying various
learning objects with different content into certain predefined classes. In this regards, an automatic classification plays a key role in organizing these massive sources of unstructured information into a structured format. Therefore, we propose a new approach to automatically organize the pedagogical multimedia content using an automatic classification based on the pedagogical course domain ontology.

The course domain ontology consists of a set of concepts in the domain of teaching and associated relations. Concepts are generally expressed through natural languages and most concepts in ontology will be represented as clusters of relevant terms. Specifically, each concept in the ontology will be formed by a list of synonym terms. This ontology development process is known as bottom-up approach ontology learning [3]. This representation is important for the use of the ontology, as it will make it easier to link the concepts in the ontology to the learning objects in actual HIP platform.

Ontology based classification approach represents semantic aspects of information coming from different media modalities through entities defined within the course domain ontology. The learning object using the domain ontology is represented as a vector where the vector components represent concepts. Concepts are extracted from ontology and their importance is calculated from the corpus using statistical measures used in traditional information retrieval tf*idf.

The contribution of ontology concepts in classification process depends on the position where they are depicted in the hierarchy and this contribution is indicated by weight. The hierarchy consists of concepts such as classes, subclasses and instances that may have different weights to represent the concepts’ importance.

Furthermore, this paper also proposes a new model to enhance the information representation by automatically estimating the weights of the concepts in the ontology. Thus, in addition to enhancing the representation model, we will improve the classification performance in terms of efficiency and effectiveness.

In the rest of this paper, an overview of previous research work with respect to implementation of course domain ontology on eLearning platforms is given, followed by a section which describes the new proposed approach. A short description of implementation is given in section 4. The paper will conclude with a short section describing the conclusions and future work.

2 Related Work

This section describes research that has been carried out by researchers in the area of domain ontologies for intelligent eLearning architectures and systems, and it examines the key difference between our work and other research.

Course domain ontology, as a knowledge for certain course domain, can be utilized at development time to clarify the meaning of the concepts and their properties for a specific domain of interest, particularly to aid understanding, communicating and facilitating system integration. This can be accomplished during the requirements engineering phase. Also, generating an eLearning course using course ontology could assist in defining a systematic approach in developing eLearning course [3], [4].