

Abstract

One of the most important components of an e-learning system is the learning material or the learning content. The popularity of e-learning has led to the development of many learning object repositories that store learning materials specifically created for e-learning. Besides, the world wide web contains many articles and good quality learning materials. High quality learning materials are expensive to create. So it is very important to ensure reuse of learning content. Reuse is made possible by annotating learning content with metadata. Manual annotation is a time consuming and expensive process. It is also liable to human errors.

In this thesis, we have worked on the automatic annotation of learning materials. We have identified a set of metadata attributes that describe some important pedagogic characteristics of learning materials. We have developed an automatic annotation tool, which annotates given learning materials and thus facilitates the creation of a learning object repository. To make the best use of the learning repository one needs to be able to retrieve learning materials that are most relevant to the learner's requirements. The metadata associated with a learning object are chosen so as to make this possible. We use as metadata pedagogic attributes like *document type*, *topic*, *difficulty level*, *coverage of concepts*, and for each concept the *significance* and the *role*. A number of methods like standard classification algorithms, parsing and analysis of documents have been used for automatic extraction of the above metadata attributes. The automatic extraction of some of the metadata makes use of the domain ontology. The domain knowledge of the subject is captured using a structural ontology of the domain and this ontology has been manually developed for a few domains.

Further, a learning system should be able to deliver personalized learning materials to a learner. To deliver personalized learning materials to a learner, we have

developed a search tool. The personalized retrieval is based on the user profile. The user profile includes what the learner already knows (the learner's knowledge state) and what the learner is required to know (the learner's curriculum requirements).

The major contribution of the thesis can be outlined as follows:

- Identification of some important pedagogic metadata attributes of learning materials to facilitate e-learning.
- Development of different algorithms for automatic extraction of the metadata attributes from the learning content.
- Development of an automatic annotation tool to facilitate the creation of a learning repository.
- Development of a search tool for personalized retrieval of learning materials.

Automatic extraction of pedagogic metadata is a sub problem of natural language processing and shares the latter's difficulties. However, because of its limited scope and the availability of the contextual knowledge in the form of ontology allows comparatively superficial analysis to give encouraging results.