1. Introduction

“When a modifiable connection between a situation and a response is made and is accompanied or followed by a satisfying state of affairs, that connection strength is increased; when made and accompanied or followed by an annoying state of affairs, its strength is decreased.”

- Edward Thorndike

This chapter presents a brief introduction of the thesis and motivation for the presented work, followed by the contributions and the organization of the chapters in rest of the thesis.

Learning through Web, internet, computer based technology or e-Learning comprises of unlimited ways to educate a learner (Nichols, 2003), (Mayes & Freitas, 2004), (Bittencourt, Costa, Soares, & Pedro, 2008). Different ways and approaches used for e-Learning may however vary depending on the learning goals or learner’s information needs. These may also vary depending on the learning approaches used by the e-Learning components themselves, such as instructional learning (Alonso, Manrique, & Vines, 2009), learner or user centered learning (Ferran, Guerrero-Roldán, Mor, & Minguillón, 2009), (Catarci, Giovanni, Gabrielli, Kimani, & Mirabella, 2008) or hybrid approaches (Rohrer & Pashler, 2010). Knowledge dissemination is a most fundamental e-Learning task that makes provisions to provide useful learning content to a learner according to a learner’s information need. Executing this task however requires a collection of meaningfully structured content so that the content can be utilized to impart required knowledge to a learner.

This collection of content is usually created with the help of experts, specifically designed for a specific curriculum (Chi Y.-L. , 2009) or embedding content in a fixed format similar to the design of Learning Objects (LO) (Koper & Olivier, 2004). However, this type of content creation which is specific to a template severely gets affected by the need of continuous efforts for its improvements and upgradations.
Introduction

It has been observed that in-spite of dedicated web portals and offline learning resources such as text-books, lecture notes and supplementary learning materials, the learners still prefer to search the WWW for a variety of the content and latest information. Needless to say, the WWW is the biggest educational content repository which gets augmented every second, collaboratively by the different communities. That is why the Web is always preferred for searching the relevant learning content. However, due to the huge volume of the Web and diverse nature of the content, there is essentially a need to create a mechanism for gathering useful learning content from the Web. Besides content gathering, there is a need to organize the gathered content to ensure an appropriate and correct association of content to the concepts for retrieval.

At present, the learners retrieve this information with the help of search engines. These search engines as a background process perform generic crawls in order to gather web documents. These documents are stored and indexed at servers. It takes a long time to crawl the complete web. Moreover, the indexes generated in this way are very large because of the generic crawling. Retrieval of relevant documents on user’s query is made from these indexes by using keyword matching (syntactic) algorithms, which brings larger number of irrelevant web documents to searchers (learners). Also, these methods of generic crawl and maintaining a global index may not work for a longer period of time due to the fast pace of increase in information on the Web.

Through the work in this thesis, a solution towards this problem is proposed. Gathering learning content from the Web and organization of content semantically can be accomplished by training software agents intelligently on individual’s need. These agents can be made to work collaboratively to find and collect useful learning content from the biggest resource repository, the WWW, structure them semantically which can be retrieved by the learner later as per his need.

1.1 Research Trends and Related Work

The literature has witnessed E-Learning systems that can be divided into two broad categories. The one enclosed within a boundary which encapsulates the learning, teaching and a support system for basic administrative activities is called
closed e-Learning environment. Primarily its motive is to facilitate the stakeholders with a virtualization of a complete university or a tutor system. In such systems, usually the learners need to register themselves in a system. The learners are accompanied by a set of human instructors/ mentors who formally help them in their educational need. A few among these systems have intelligent features like adaptiveness and personalization (Brusilovsky, 1999). Such systems somehow did not get much popularity among the community, perhaps due to the complicated authoring procedures and the expertise required for creating and updating the knowledge base. Though the closed environment e-Learning systems that have evolved in the last decade incorporate intelligent features, they are usually inclined towards the curriculum based framework. A few prominent systems in this category include Intelligent Tutoring System (ITS) (Nwana, 1990), Interactive Learning System (ILS) and, Adaptive and Intelligent Educational System (AIES) (Brusilovsky, 1999). Systems such as ITS and ILS focus on problem solving support technology, curriculum sequencing, intelligent analysis of student’s solutions and interactive problem solving support (Devedzic, 2004). AIES inherits intelligent techniques from ITS and Adaptive Hypermedia Systems (AHS). These systems are adaptive and intelligent web based educational systems that majorly focused on the course sequencing for a student according to his/her needs. This requires a definite set of resources (in the form of lessons) and various patterns (sequence in which these lessons may be arranged) which are used as database for performing intelligent tasks. Such systems are generally more suitable for the pedagogy strategies of learning theory, but at the same time require lot of efforts in the customization of a system. It requires constant up-gradation of lecture notes, lessons, activities and development of various learning patterns according to different categories of learners.

Andrology strategies of learning theory (Nichols, 2003) which are more applicable in higher education, to adult learners as they are more oriented towards self-directed learning (Merriam, 2001), need lots of educational resources (learning material). Here, learners are not only students, but they may also be instructors, teachers or educators. Learning, in reality is not limited to a subject or curriculum. It may get hampered if a learner is confined to limited resources and areas (Clark &
Mayer, 2011). Any course or curriculum based learning usually has defined boundaries and limitations regarding the depth of the content and the level of information required by a learner, therefore, a learner may not get an opportunity of a diverse range of learning resource material exposure. However, matured learners, who believe in self-directed learning usually, refer all other possible resources to understand the concept in its completeness. This is also evident from the research by Head & Eisenberg (2009) which shows that the learners in higher education often seek possible online resources for their information needs related to academics. Moreover, learning is not an isolated process rather it happens in parallel with the regular day-to-day activities of students and educators (Jovanović, Gašević, Torniai, Bateman, & Hatala, 2009).

These concerns over learning focus the attention towards an equally important area as the other category of e-Learning systems i.e., open e-Learning system. The open e-Learning system is an open i.e., accessible to all, and Social Semantic Web based educational system which is still in its developing stage and is an ongoing research field. It is not confined to the closed boundaries (therefore it is open) in terms of learning, i.e., neither it is customized to any organization/university nor it works only within a limited set of learners/educators. It has lot of similarities to the other web based application where anyone can store/ retrieve information from the dynamically updated web, either using web services (Lin, Holt, & Li, 2006) or developing software agents to perform designated tasks.

The research community in this area has put efforts to make available an Open Educational Learning system with the intent to cross the boundaries of predefined and fixed curriculum based learning. A lot of features of systems such as EdShare (2008), Connexions (2009) and MERLOT¹ inherit the characteristics of open e-Learning systems. They utilize semantic techniques such as semantic web, knowledge organization, reasoning, intelligent software agents for the benefit of open e-Learning. The system proposed in the thesis also belongs to this category. These systems work towards the creation of an enriched environment for self-learning and thus providing

¹ http://www.merlot.org/merlot/index.htm
an opportunity to learner for achievements in higher education. A few of such research/systems have been discussed below.

MERLOT (Multimedia Educational Resource for Learning and Online Teaching) is a project created by the California State University Center for Distributed in 1997. It has over 83,000 members who share and reuse open resource educational repository (Davis, 2011). It contains a collection of peer-reviewed educational materials pertaining to world languages in its repository. It is an initiative towards a collective advancement of education through open technology, open content and open knowledge (Hanley & Carey, 2004).

Roy (2006) has proposed an approach in her work for automatic annotation of learning materials with metadata in order to create e-Learning annotated repository. This may be used by different learning management systems or tutoring systems. The author has claimed to develop a search tool for retrieval of the appropriate learning materials to meet a learner’s requirements. The retrieval module searches these documents from the repository and also from the web. These documents are further analyzed for the learner’s relevance using domain ontology before presenting them to the learner. From pedagogical point of view, the documents have been classified only in three types from the large set provided in the IEEE LOM (Learning Object Metadata) resource type, for which some semantic preprocessing is required to achieve acceptable results (Roy, Sarkar, & Ghose, 2008).

OER (Open Educational Resources) has emerged as a concept with great potential to support educational transformation (Butcher, 2011). The Commonwealth of Learning (COL) is an intergovernmental organization created by Commonwealth Heads of Government to encourage the development and sharing of open learning/distance education knowledge, resources and technologies. COL primarily shapes and promotes new models for OER generation, publication, access and re-use.

In India, a few recent attempts have started OER experimentation, which includes, i) the National Programme on Technology Enhanced Learning (NPTEL) that provides digitized and web-based lectures on engineering courses to faculty and

---

2 http://www.col.org/Pages/default.aspx
3 http://www.col.org/progServ/policy/Pages/oer.aspx
students; ii) The Virtual Academy for the Semi-Arid Tropics (VASAT) that provides non-formal “natural resource literacy” to rural communities using simple technology tools; and iii) FlexiLearn, an initiative of the Indira Gandhi National Open University (IGNOU) that expands the scope of open and distance learning by providing “free learning resources integrated with a Learning Management System” to enhance personal learning free of cost (Harishankar, 2012).

The system presented in the thesis is different from all the above mentioned e-Learning systems or OERs in the design and methods used in its working. The proposed system has been designed as a multi agent system and uses semantic web technologies. It automatically creates and updates the content repository and maintains the relationship between concepts and content as well. Crawler agents built as a part of the system collect contents from the Web and social portals automatically. There are other agents in the system for managing the knowledge base and imparting this knowledge to a learner. Thus, these agents collaboratively perform the important tasks of content retrieval from the Web, organizing them in the knowledge base and responding to a learner’s needs by utilizing this knowledge base.

1.2 Motivation

Mentoring is a process of knowledge dissemination by a more knowledgeable person (called Mentor) to a less knowledgeable person (called Mentee). Mentoring is referred as informal when the conversation between mentor and mentee takes place spontaneously, usually for a short duration, till the immediate goal of the mentee is achieved. Also, it does not require any prior investigations on the mentor-mentee relationship, in case of mentoring being informal.

The proposed work has been influenced by the needs and the benefits of informal mentoring which a student in a role of a learner or a mentee receives in his/her life, throughout the academic career from various sources. Some of these sources are depicted in Figure 1-1. Every source of information has its own importance and advantages which guides/assists a mentee in one or the other way, such as concept understanding, getting related information about a concept for astute analysis, preparing notes for assessments and examinations etc.
It has been found that besides receiving sufficient knowledge from instructors, a mentee (learner) still approaches senior and/or peer students for some additional information related to the subject. Not limiting themselves to this information, mentees usually search various web sites for some additional examples and many times surf the social sites to gather communities’ opinion, problems or solutions about the subject/ concept of their academic interest. However, in the whole process of gathering information of interest and need, lot of mentee’s time gets consumed which hampers self-learning. Alternatively, if through some mechanism the information of interest to a mentee gets retrieved automatically from all these sources, then it would benefit the mentee in saving the surfing time and the time consumed in visiting individual mentors. This would effectively improve and enhance the learning experience of a mentee.

At the other hand, the future prospects of the Semantic Web fascinated and motivated us to construct a semantic approach for mentoring tasks by simplifying the retrieval and search of various useful web resources. Today the Web is overwhelmed with information. A person spends more time in searching the required content than actually understanding it, which consequently slow down the actual learning pace. The work in this

Figure 1-1: Informal mentoring sources for a mentee in a real scenario

At the other hand, the future prospects of the Semantic Web fascinated and motivated us to construct a semantic approach for mentoring tasks by simplifying the retrieval and search of various useful web resources. Today the Web is overwhelmed with information. A person spends more time in searching the required content than actually understanding it, which consequently slow down the actual learning pace. The work in this
thesis is thus an attempt to deploy software agents for this searching on behalf of a learner. The next section describes the contributions and organization of the rest of the thesis.

1.3 Contributions and Organization of Thesis

This thesis with the intent to provide solution to open e-Learning, aims to present a design for educational concepts using RDFS/OWL semantics. The semantics are particularly useful to expand a concept for retrieving and searching semantically relevant information on the Web. Besides the ontological designing, efforts have been put to build crawlers that are capable of retrieving potentially relevant web resources from the World Wide Web and Social Bookmarking Sites. These retrieved web resources are further organized according to the relevance of each individual concept in the ontology. In order to rank the potentially relevant retrieved web resources, a ranking algorithm has also been proposed. The top ranked web resources above a threshold value are considered for augmentation in the ontology. An automatic ontology augmentation approach has been proposed to organize these web resources in the ontology by semantically linking the content to various concepts with appropriate annotation.

Figure 1-2: Thesis contributions and organization
An informal mentoring system, based on a multi agent framework has been proposed to serve mentee’s information needs by knowledge dissemination. This system makes use of the Concept Ontology. The proposed system incorporates the communication among agents to achieve the collaborative efforts for the resource retrieval task through Focused Crawlers and finding their relevance to various concepts, the task of automatic organization of resources in ontology and finally providing potentially useful information to a learner.

The resultant developed system would benefit mentor and mentee both, as the mentee would get relevant information with all-time availability. (S)he would not need to surf various search engines by trying different combinations of query terms, and neither need to visit human mentor for his/her academic related information needs. The mentor in such a scenario would able to get time for other mentoring activities.

Figure 1-2 illustrates various components of the proposed system. It shows the major contributions of the thesis and the chapters to which each contribution is associated. Although all components handled by agents work asynchronously, but the process of Information Retrieval and organization of retrieved web resources which is a continuous background process follows a sequence consisting of these components as sub-processes. The encircled numbers in the figure represent this sequence. An attempt has been made to arrange chapters in the thesis according to this flow. Chapter 1 introduces the basics of the e-Learning system followed by the research trends and motivation to the thesis work. Rest of the thesis is organized as follows:

- **Chapter 2** discusses the background of the Web and its different versions. It also presents the basic concepts, tools and techniques used throughout the thesis work. The agent paradigm, its tools for implementation and the need to use them are also discussed in this chapter.

- **Chapter 3** demonstrates the modeling of educational concepts and content in Ontology and presents an approach to automatically organize retrieved resources in the knowledge base. Term Expansion, one of the important tasks in Information Retrieval has also been described in the chapter, as it takes
advantage from the ontology for the expansion procedure.

- **Chapter 4** primarily brings up the design of two types Focused Crawlers that retrieve potentially relevant learning content from different types of web repositories. The prototypes FCHC (Focused Crawler based on Human Cognition) that explores Social Bookmarking Site for the useful content and, DSRbasedSFC (Dynamic Semantic Relevance based Semantic Focused Crawler) that crawls on the WWW are also demonstrated in the Chapter.

- **Chapter 5** focuses on the methods for ranking web pages based on a topic’s relevance. The chapter also discusses the proposed algorithms:
  - Content based Semantic Ranking (CSR) that ranks untagged web resources using Semantic Term Frequency (STF) and,
  - Social Semantic Ranking (SSR) that ranks tagged web resources using Social Semantic Relevance (S2R), based on Vector Space Model.

- **Chapter 6** demonstrates the design and implementation of a Multi Agent System for e-Mentoring, a sub-task of e-Learning. The framework primarily incorporates a Virtual Mentor who intelligently processes a mentee’s information needs, retrieves relevant information from other co-operative semantic information agents for delivering the relevant content to a mentee. The Chapter also presents a detailed experimental study and evaluation of the complete work pertaining to astute analysis.

- **Chapter 7** finally concludes the thesis summarizing the contributions and gives some suggestions for future work.