Available online at www.sciencedirect.com
ScienceDirect
Applied Soft Computing 7 (2007) 746–771
Applied Soft Computing
www.elsevier.com/locate/asoc

MASACAD: A multi-agent approach to information customization for the purpose of academic advising of students

Mohamed Salah Hamdi

University of Qatar, College of Engineering, Department of Computer Science, P.O. Box 2713, Doha, Qatar

Received 11 March 2004; received in revised form 19 January 2006; accepted 23 February 2006
Available online 18 April 2006

Abstract

The growth and advancement in the Internet and the World Wide Web has led to an explosion in the amount of available information. This staggering amount of information has made it extremely difficult for users to locate and retrieve information that is actually relevant to their task at hand. Dealing with this problem of "information overload" will need tools to customize the information space. In this paper we present MASACAD, a multi-agent system that learns to advise students by mining the Web and discuss important problems in relationship to information customization systems and smooth the way for possible solutions. The main idea is to approach information customization using a multi-agent paradigm in combination with a number of aspects from the domains of machine learning, user modeling, and Web mining.

Keywords: Information customization; Multi-agents; E-Learning; Machine learning; Neural networks; User modeling; Web mining; Expert systems

1. Introduction

The recent proliferation of personal computers and communication networks has a strong scientific, intellectual and social impact on the society. Rapidly evolving network and computer technology, coupled with the exponential growth of the services and information available on the Internet, has already brought us to the point where hundreds of millions of people should have fast, pervasive access to a phenomenal amount of information, through desktop machines at work, school and home, through televisions, phones, pagers, and car dashboards, from anywhere and everywhere. The challenge of complex environments is therefore obvious: software is expected to do more in more situations, there are a variety of users, there are a variety of systems, there are a variety of interactions, and there are a variety of resources and goals.

To cope with such environments, the promise of information customization systems is becoming highly attractive.

The recent popularity of the World Wide Web (Web) has provided a tremendous opportunity to expedite the dispersement of various information creation/diffusion infrastructures. The mass of content available on the Web raises important questions over its effective use. With largely unstructured pages authored by a massive range of people on a diverse range of topics, simple browsing has given way to filtering as the practical way to manage Web-based information. Today's online resources are therefore mainly accessible via a panoply of primitive but popular information services such as search engines.

Search engines are very effective at filtering pages that match explicit queries. Unfortunately, most people find articulating what they want extremely difficult, especially if forced to use a limited vocabulary such as keywords. The result is large lists of search results that contain a handful of useful pages, defeating the purpose of filtering in the first place. Search engines also require massive memory resources (to store an index of the Web) and tremendous network bandwidth (to create and continually refresh the index). These systems receive millions of queries per day, and as a result, the CPU cycles devoted to satisfying each individual query are sharply curtailed. There is no time for intelligence. Furthermore, each query is independent of the previous one and no attempt is made to customize the responses to a particular individual.

What is needed are systems that act on the user's behalf and that can rely on existing information services that do the resource-intensive part of the work. These systems will be sufficiently lightweight to run on an average PC and serve as personal assistants. Since such an assistant has relatively modest resource requirements it can reside on an individual

E-mail address: mhamdi@qu.edu.qa.

1568-4946/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2006.02.001

user's machine, which facilitates customization to that individual. Furthermore, if the assistant resides on the user's machine, there is no need to turn down intelligence. The system can have substantial local intelligence and information customization becomes possible.

The work described here discusses some ideas aiming at improving the reduction of information overflow, which is so common today in many domains. The general framework is that of doing information customization based on an multi-agent approach.

In previous work [1–5], we presented an E-Learning system that provides a service to a student that checks whether lecturers are offering information that matches the profile of the student and informs the student of the information found. The student is registered in many courses and seeks course materials from the sites of the different lecturers. These diverse resources can be physically distributed. They are also dynamic so that course materials can be added, updated or deleted. The student profile, that includes the courses attended currently and possibly more information, changes also over time because the student can leave a course or register in a new one. This means that the customized presentation of information for the student should be updated continuously as new information becomes available. This happens with no user intervention using an autonomous multi-agent system.

In this paper, following the same long-term objective of providing a complete E-Learning environment for students and striking for the more general goal of information customization, we describe MASACAD (Multi-Agent System for ACademic ADvising; ''MASACAD'' is also the Arabic word for ''courses''), a multi-agent system that advises students by adopting a machine learning paradigm. Machine learning methods can be used to deal with many different aspects of the problem of advising.

Academic advising, in its simplified version, consists of telling the student which courses he/she should register in based on the profile of the student, on the university laws, and on the courses that are offered in the semester for which advising is needed.

In the following we will discuss some important points in relationship with E-Learning, information customization, user modeling, agent systems, machine learning, and Web mining. Ideas from all these domains are combined to devise a solution for the problem of academic advising of students. After that, we will then describe the details of the learning multi-agent application.

2. E-Learning

E-Learning is a valuable extension of the distance education paraphernalia, enabled by the new information and communication technologies. Distance education normally occurs in a different place from teaching and as a result requires special techniques of course design, special instructional techniques, special methods of communication, as well as special organizational and administrative arrangements [6]. E-Learning is often described as the use of network technology, namely the Internet, to design, deliver, select, administer and extend learning.

Due to the flexibility provided to students and teachers, both in space and time, E-Learning may be a source of great joy to its users and an important source of financial resources for many organizations. E-Learning is based on the cooperation of geographically distributed participants, and many of the activities the participants are supposed to perform do not have strict time schedules, but do have time constraints that must be respected. If these constraints are not fulfilled, severe problems may occur and the success of a specific task or action may be in jeopardy. These kinds of problems are very difficult to handle and solve, because of the distributed nature of the resources and participants of an E-Learning environment.

One key issue in E-Learning is communication between participants, for which there are two basic types of technological solutions: asynchronous and synchronous. In the asynchronous approach, the interaction between parties does not require them to be engaged at the same point in time. In synchronous communications the interaction between participants requires simultaneous engagement of the participants. Examples of technologies for asynchronous communications are hypertext publication (namely www), email, mailing lists, newsgroups/bulletin boards and file download (ftp). For synchronous communications the more often used technologies are: chat/IRC, whiteboard, audio/video streaming and videoconference. The existing E-Learning platforms, such as WebCT, Lotus Learning Space, Blackboard, Centra, etc., incorporate both models, and corresponding services, in different ways.

Online education is today a reality in many sectors of the society, especially in educational centers such as colleges and universities, increasingly high schools, and also professional groups demanding continuous access to education. Several years ago this new educational method was considered as an experimental approach with more disadvantages than advantages. However, today it should be considered not only a complementary educational resource but also a serious alternative that competes to conventional and now classical methods. Both methods will coexist and the logical initial inertia to ignore the new opportunities provided by the new media should be reduced and be faced sooner better than later in the same manner in which many other areas were modified throughout history.

Obviously the adaptation to the new features and services of the E-Learning environment is not immediate and requires experience, time, investment, pedagogical and technical resources, and government or campus administration support. At the UAE University there exits enormous interest in the area of online education. Rigorous steps are taken towards the creation of the technological infrastructure (hardware, software, and communications) and the academic infrastructure (course materials, teacher–student communication) for the improvement of teaching and learning. MASACAD, the academic advising system described in this paper is to be understood as a tool that uses network technology to support learning and as part of the E-Learning environment at the