National University of Ireland, Galway
Ollscoil na hÉireann, Gaillimh

Department of Information Technology

technical report NUIG-IT-220901

Bridging the Gap between Desirable Web-based Education Systems and Current Technical Solutions

P. Byrne (NUI, Galway)
S. Flynn (NUI, Galway)
J. Griffith (NUI, Galway)
C. O'Riordan (NUI, Galway)

Bridging the Gap between Desirable Web-based Education Systems and Current Technical Solutions

Pat Byrne, Sharon Flynn, Josephine Griffith, Colm O' Riordan
Department of Information Technology,
NUI, Galway.
{pat.byrne, sharon.flynn, josephine.griffith}@nuigalway.ie, colmor@geminga.nuigalway.ie

Abstract

The increased use of the web as a means of disseminating material seems a dominant trend in third level institutions. This modern development is the latest in a long trend in harnessing technology to mimic the traditional classroom model. However, systems to-date do not deal with many of the issues recognised in pedagogical research and instead merely present information via a new medium. Techniques from AI seem to offer the potential to address some of the oft overlooked evidence from the pedagogy research community.

This paper discuss some topics from the arena of education research and their incorporation into automated educational systems. A review of many of the existing approaches is presented. The paper also outlines ideas from AI which offer hope in bridging the gap between the desirable requirements of web-based education systems and current systems.

1 Introduction

This paper outlines current wisdom in education and automated educational systems, in particular, web-based education. We provide a critique of many of the existing approaches in the domain. We propose techniques from the field of Artificial Intelligence as potential contributions to the development and deployment of web-based education systems. Section 2 discusses some issues with respect to effective, quality teaching and education. The next section discusses existing computer-aided approaches including tutoring systems and web-based systems. Section 4 outlines ideas within AI which may help bridge the gap between the desirable requirements of a computer-aided education system and current systems.

2 Education

The practice of lecture-based teaching has been repeatedly questioned[10],[4]. Teaching by lecture is often a default choice partly chosen because it is familiar, gives control to the lecturer, and is economically viable.

With the increased interest and attention payed to "anytime, anywhere" learning, the traditional methods of teaching are being further questioned, with an upsurge in research into exploring alternative viable means. Much attention, in particular, is focussed on the well-established concept of lifelong learning—it is accepted that what we learn through the traditional educational cycle no longer equips us with all the knowledge which we will need throughout a working life.

This provides educationalists with new challenges to meet a changing and varied group of students, of different ages and with different abilities and experience. Trying to maximise the learning experience for each student requires a flexible learning environment, allowing them to build on existing abilities.

Current pedagogical thinking recommends student-centered approaches to learning leading to an emphasis on personalised learning, evaluation, personalised tailored feedback, and collaboration among students.

Educational establishments are also trying to come to terms with providing a plethora of new courses within a limited budget, and rationalisation decisions often mean that the student is placed in large classes with less opportunity to engage with fellow students or receive individual feedback and support.

Over the past years there has been an increase in the use of computers and the world wide web (WWW) to aid in the delivery and management of instructional material. Web solutions provide the opportunity to develop a flexible environment and also can facilitate a large range of students in different locations, providing education when and where it is required by the student. In contrast to traditional lectures, a web-based education system may encourage more active participation as more control and responsibility is given to the student (e.g. order of topics, pace). Therefore, the WWW offers many advantages such as world-wide delivery, access to a huge amount of information, an attractive media for representing course material, support for a non-linear approach to learning, and no time constraints. In addition, over time, links can be tailored to suit individual students.

Having these new modes of delivery is not simply lecture delivery through another medium. Using the technology appropriately we can also provide a motivational and exciting multimedia environment which will help to maximise the learning experience.

Existing systems do not necessarily provide a classroom environment—the experience of being part of a learning community which can generate synergy and provide a collaborative base in which educational growth will flourish is often missing.

In summary, the desirable requirements of any such computer aided education system should provide for "anywhere anytime" learning to cater for recognised varied student profile, should be designed to be pedagogically sound, should provide appropriate tailored feedback, and should also be cost effective.

3 Tutoring Systems and education

3.1 Early Computer Aided Instruction

The earliest Computer Aided Instruction(CAI) systems were developed as a direct result of research carried out by behavioural psychologists on *programmed learning*. The basic strategy is that material is presented in small steps and the student is required to answer questions. The user is provided with immediate feedback on their responses. The number of correct and incorrect responses produced are recorded by the system and may be evaluated by the teacher and the student at a later time. In a true Skinnerian system, the presentation of material and questions occurs in a fixed sequence.

However, this linear approach to learning does not even attempt to accomplish the concept of *individualised learning*, where each tutoring session is directed at the particular student involved in it. This has always been one of the more predominant goals of computerised tutoring[9].

One way of attempting to achieve such individualisation is in the use of branching programs where the next question asked depends entirely upon the answer given to the previous question. Each student will, then, follow a different path depending on the answers he gives to the questions. Another approach to the problem is to have a number of difficulty levels built in. The system should then set itself to the appropriate level depending, again, on the answers given by the student. In neither of these approaches does the system try to *understand* the student to any degree.

The general strategy of CAI has been regarded as being too restrictive. At a time when education theory was moving on, the use of CAI risked reversing the process, advocating the use of multiple-choice questions rather than opening up possibilities for exploration for the learner[6].

3.2 Intelligent Tutoring Systems

In order to address some of the shortcomings of CAI, in particular that of individualisation, attempts were made in the 1960's to apply some of the techniques of Artificial Intelligence (AI) to computerised tutoring systems. The main difference between CAI and ITS is that instead of attempting to encode the decisions of experts in the form of programs, the field of ITS concentrates on attempting to capture the very knowledge that allows experts to compose an instructional interaction in the first place. The programs themselves make the decisions based on the knowledge with which they have been provided. This introduces the possibility that the systems could make decisions not anticipated by the experts.

The research into the application of AI techniques to computerised tutoring soon abandoned one of CAI's early objectives – that of providing total courses. The reason for this is that while CAI has primarily been developed by those interested in its application to education, ITS has, for the most part, been the domain of researchers in computer science. These AI researchers are only too aware of the limitations of what can be done. "Given that we understand little about what constitutes good teaching, building a good teaching system presents certain difficulties" (Ennals[6]).

Although there is no standard approach to the development of ITS, most researchers agree to the division of such development into four distinct components[14]. These comprise: domain expertise; student model; communication strategies or pedagogical expertise (referred to as the pedagogue); and interface with the student.

The expert module is that component of an ITS which stores some representation of the domain to be tutored. The module should have the ability, to some extent, to manipulate this knowledge so as to give the impression of *understanding* and being able to reason with the knowledge. These manipulations may or may not be similar to the process which humans use for reasoning about the domain.

The pedagogical component of an ITS is intended to contain knowledge about *how* to teach. The role of the pedagogue can vary from one ITS to another. Decisions are made by reference to the student model and to the expert module. The pedagogue then adapts its presentation of topics to the needs of the individual students. The module can also determine when to interrupt the student, which questions should be asked and in what order, when to give explanations, what could and should be said or presented at any given time. This can provide the desired flexibility in such systems. It may also be the role of the pedagogue to examine the student's answers, with reference to the expert module, and thus cause the student model to be adjusted.

In order to provide individualised instruction, the system must be provided with some representation of, or some way of representing, the student. Ideally the student model should contain all the information about the student which is relevant to the teaching process; current knowledge state of student, behavioural tendencies or the student, attitudes of the student and many other aspects of the student play a part in the education process. Techniques for user modelling are presented in section 4.2.

The user interface is the only channel of communication between the student and the ITS. The interface can make the presentation of information understandable to the student, and this can be a crucial factor in whether or not the student will learn from the ITS.

3.3 Web-Based Teaching and Learning

Ausserhofer[2] identifies a number of categories of computer supported learning environments including: Online Books; Edutainment Software; Online Education.

Online Books This involves the preparation of electronic versions of lecture notes and books, and is widespread in third level education. From a lecturer's point of view, it is the fastest and easiest approach, and it dramatically reduces the amount of photocopying required for large classes. However, it is the worst and most inefficient approach with respect to educational success.

Edutainment Systems These are systems based on the observation that children are more interested in computer games than in educational software. The game model uses levels of difficulty in an easy and fun progression. However, the structure of these education programs have to follow a storyboard in order to keep up interest, and the goals of the game must be closely related to the educational material under tuition. Thus it is quite complicated to apply a game model for educational purposes, and it is not clear that children actually learn just because they enjoy playing a game.

Online Education – Tele-Teaching and Learning The advent of high power communication technology now means that video cameras and audio facilities can be used to transmit live lectures online in synchronous mode, allowing students to take part remotely. Students can even interact with each other using high-speed communication facilities.

Online Education – Web-Based Teaching and Learning These systems are asynchronous, in that teachers and students are not required to be present simultaneously. In general they do not require any sophisticated equipment, unless combined with other approaches.

In the last number of years, the Internet has opened up enormous possibilities for computer supported learning environments. However, in transferring their material to the web, educators seem in danger of returning to the old model of *education as transmission*. It is common to talk about *delivery* of course content, as if a student is some sort of receptacle for knowledge. On the other hand, researchers in internet technologies use web-based education as an easy and obvious application, without questioning if the technologies actually help or hinder learning.

The use of the web for education should support the role of ITS in providing individualised instruction. In addition, the advantages of the internet can be used to further enrich the learning experience, through the vast amount of information available, and by enabling student-to-teacher and student-to-student communication.

Two approaches which have been used with some success to facilitate peer learning include the use of online discussion boards[11, 7] and web-based peer review[8].

Some of the more successful web-based education environments combine a number of approaches, including online lecture notes, transmission of live lectures, the interactive classroom, online quizzes and assessments, and simulations[13]. In many cases this is in addition to traditional instruction.

4 Artificial Intelligence and Education Systems

4.1 Introduction

In any useful education system, be it a tutorial system or a web-based education system, it appears necessary that to make a reasonable effort at overcoming shortcomings inherent in traditional computer based learning packages, that some degree of intelligence should be incorporated in the system. Some possible approaches are outlined in the following sections; they include user-modelling to move towards providing personalised learning, collaborative filtering which aims to recommend suitable material to people and goes some way to simulating a classroom environment, albeit in a limited way and finally intelligent organisation of material.

4.2 User modelling

The aim of a user modelling system is to learn about a user or group of users. A user modelling system must record the user's actions and statements to form a user model. This information must then be processed to transform the data in to a representation useful for the system. User modelling is often difficult because:

1. partial results are often needed before there is opportunity to learn enough about the user; these partial results are often uncertain, so user modelling techniques must be designed to handle this.

2. users change over time leading to problems with the consistency of a model. For example, user preferences may change; users may forget things; users may adopt new goals on a whim; users may abandon previous goals, etc.

3. some current techniques are computationally expensive

User model techniques can be characterised with respect to the following features: passive/active gathering of user preferences; whether the gathering of such data is user-initiated or automatic and effected in a direct or indirect manner; and finally whether the construction of such a model occurs on-line or not.

In a web-based education context, it is of the utmost importance to maintain user models so as to try to provide personalised information. This model can be built from explicit evidence provided by potential users—background and experience in the given topic etc., and from implicit evidence learned from monitoring a student's behaviour—results in exams, pages visited and revisited (and their associated difficulty) etc. It is possible to build large user models but difficulties outlined above do exist in building accurate models and in the correct deployment of such.

4.3 Collaborative Filtering

Collaborative Filtering is not based on analysis of the content of the document set but on the premise that "people with similar interests in the past will have the same interests and preferences in the future"[12]. Collaborative filtering systems and recommender systems attempt to exploit this information to predict users' interests.

Given a set of users, a set of items, and a set of ratings, systems attempt to recommend items to users based on prior ratings. The collaborative filtering systems essentially automates the "word of mouth" process.

Collaborative filtering offers a number of advantages over content-based filtering. The most obvious difference between the two types of filtering is:

- content filtering fails to capitalise on the knowledge and opinions of people who have previously accessed documents in the document set.

- with collaborative filtering, attributes such as quality, clarity, presentation style, and not just content, can be taken into account.

The problem space can be viewed as a matrix consisting of the ratings of each user for the items in the document set, i.e., the matrix consists of a set of ratings $u_{i,j}$, corresponding to the rating by user i for an item j. Using this matrix, the aim of collaborative filtering is to predict the ratings of a particular user, i, for one or more items in the document set.

Neighbourhood-based algorithms are the most commonly used approach in collaborative filtering. A subset of users is chosen based on their similarity with a current or active user. Such methods comprise three main steps:

1. select a set of users with similar interests/preferences to user i, i.e., users who have similar ratings for items as user i, calculate user correlation.

2. select a subset of these users as a set of predictors

3. predict recommendations for user i from the set selected in step 1, i.e., if these users rated an item j highly, this item will be recommended to user i

Various techniques can be used to calculate the user correlation. These include Pearson correlation (a weighted average of deviations from the neighbours' mean is calculated), Constrained Pearson correlation (a variation on Pearson correlation where the deviation from the median of available rating values), the Spearman rank correlation (uses ranking as opposed to explicit rating values, thus giving greater independence of the range of ratings), Vector similarity (uses the cosine measure between the user vectors to calculate correlation) and Mean-square difference algorithm (the mean square difference between each pair of users is calculated). Given these correlation values, a set of suitable neighbours is selected. The weighted mean of these neighbours' ratings can be used to generate a prediction.

The basic premise behind collaborative filtering mirrors, albeit in a limited and simplistic way, the means in which students will recommend material and notes to fellow students. In the domain of web-educational systems, it is possible to cluster users based on similar learning styles based on the similarity of user models. By offering suggestions to students based on these clusters we can possibly improve the quality of the learning experience and move some way towards reducing the sense of isolation often felt in distance learning environs.

4.4 Information Management

With the increased use of computers for the storage and management of large volumes of largely unstructured information, the need for automatic computerised techniques to aid users has become an issue of critical importance. A large set of methods have been developed to help users in separating relevant material from irrelevant material (information retrieval and filtering).

Information retrieval (IR) is a well established field in information science, which addresses the problems associated with retrieval of documents from a collection in response to user queries. The chief mechanisms arising in IR systems are as follows: representation (of the user's information need and the document set), comparison (of profile representation and document representation), and feedback (to improve the performance by allowing the user state his/her satisfaction or dissatisfaction with returned documents).

A range of models and techniques exist ranging from simple string matching techniques augmented by Boolean and proximity operators to more expressive and effective model such as the vector-space model[1] (in which documents and queries are represented as vectors of weighted terms and similarity is measured as the cosine of the angle between the vectors), latent semantic indexing (another vector based approach[5], which attempts to account for latent relationships between terms, has been shown to outperform the vector space model), probabilistic models[3], which model the information retrieval process in a probabilistic framework.

The application of intelligent information management in the domain of web-based education has many facets ranging from simply supporting intelligent search mechanisms (to allow students access relevant material quickly and easily) to more advanced approaches where material is clustered according to different themes depending on the student's current focus of interest. IR approaches can also be augmented to allow students easily classify and categorise large volumes of related material.

5 Conclusion

The increased use of the web as a means of disseminating material seems a dominant trend in third level institutions. This modern development is the latest in a long trend in harnessing technology to mimic the traditional classroom model. However, systems to-date do not deal with many of the issues recognised in pedagogical research and instead merely present information via a new medium. Techniques from AI seem to offer the potential to address some of the oft overlooked evidence from the pedagogy research community.

This paper discuss some topics from the arena of education research and their incorporation into automated educational systems. A review of many of the existing approaches was presented. The paper also outlines ideas from AI which offer hope in bridging the gap between the desirable requirements of web-based education systems and current systems.

References

[1] S. G. A. and M. J. McGill. Introduction to Modern Information Retrieval. 1983.

[2] A. Ausserhofer. Web-Based Teaching and Learning: A Panacea? *IEEE Communications Magazine*, 1999.

[3] J. P. Callan, W. B. Croft, and S. M. Harding. The INQUERY Retrieval System. Technical report, Department of Computer Science, University of Massachusetts, 1993.

[4] L. B. Curzon. *Teaching in Further Education: An Outline of Principles and Practice*. Cassel Education Limited, 4 edition, 1990.

[5] S. Dumais. Latent Semantic Indexing(LSI) Routing for TREC-3. *TREC 3 Proceedings*, November 1994.

[6] R. Ennals. Overview of Computer Assisted Learning. *Journal of Artificial Intelligence in Education*, 2(2):5–16, 1991.

[7] J. Griffin. Using a Web Based Collaborative Learning Management Tool to Teach Professional Issues. In *2nd Annual Conference of the LTSN Centre for Information and Computer Sciences*, London, 2001.

[8] E. Z.-F. Liu, S. J. Lin, C.-H. Chiu, and S.-M. Yuan. Web-Based Peer Review: The Learner as both Adapter and Reviewer. *IEEE Transactions on Education*, 44(3):246–251, 2001.

[9] N. D. Livergood. From Computer Assisted Instruction to Intelligent Tutoring Machines. *Journal of Artificial Intelligence in Education*, 2(3):39–50, 1991.

[10] B. R. and J. Hartley. *Teaching and Learning in Higher Education*. Harper & Row, 1984.

[11] J. Read. Using Groupwise to Stimulated Academic Discussion with Computing Students. In *2nd Annual Conference of the LTSN Centre for Information and Computer Sciences*, London, 2001.

[12] U. Shardanand and P. Maes. Social Information Filtering: Algorithms for Automating "Word of Mouth". *Computer-Human Interfaces (CHI '95)*, 1995.

[13] K. J. Siddiqui and J. A. Zubairi. Distance Learning using Web-Based Multimedia Environment. In *IEEE Academia/Industry Working Conference on Research Challenges*, 2000.

[14] E. Wenger. *Artificial Intelligence and Tutoring Systems*. Morgan Kaufmann, 1987.