The Ecological Approach to the Design of E-Learning Environments: Purpose-based Capture and Use of Information About Learners

Gord McCalla

Abstract:

The semantic web movement has grown around the need to add semantics to the web in order to make it more usable by people and by information systems. In this paper I argue that even more important than semantics is pragmatics; that is, to really enhance web usability it is critical to capture and react to aspects of the end use context. Most centrally, to make the web truly responsive to human needs, we need to understand the "users" of the web and their purposes for using it. In this paper I elaborate this argument in the context of of e-learning systems. I propose an approach to the design of e-learning systems that I call the *ecological approach*. Moving from the open web to repositories of learning objects, I show how the ecological approach shows promise not only to allow information about learners actual interactions with learning objects to be naturally captured but also to allow it to be used in a multitude of ways to support learners and teachers in achieving their goals. In a phrase, the approach involves attaching models of learners to the learning objects they interact with, and then mining these models for patterns that are useful for various purposes. The ecological approach turns out to be highly suited to e-learning applications. It also has interesting implications for e-learning research, and perhaps even for research directions for semantic web research.

Keywords: ecological approach, semantic web, learning object repositories, learner modeling, collaborative filtering, data mining, clustering

Commentaries:

All JIME articles are published with links to a commentaries area, which includes part of the article's original review debate. Readers are invited to make use of this resource, and to add their own commentaries. The authors, reviewers, and anyone else who has 'subscribed' to this article via the website will receive e-mail copies of your postings.

Gord McCalla, *ARIES Laboratory, Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5A9, CANADA, mccalla@cs.usask.ca*

1.   Introduction

The promise of e-learning goes back to the 60's, with the early computer-assisted instruction systems. Subsequent research and development has led to a plethora of computer-based paradigms: intelligent tutoring systems, collaborative learning, educational multi-media, situated learning, interactive learning environments, computer assisted learning, discovery learning, etc. There are vigorous, often fierce, debates among proponents of various of these paradigms, debates that sometimes lead to insights, but more often are self-serving and counterproductive. The debates are often framed by the different backgrounds of the proponents: whether they are from education, computer science, engineering, anthropology, sociology, psychology, library science, commerce, etc. It is clear that there are insights to be gained from virtually all perspectives on e-learning, and that to make progress we need to draw on all of these perspectives.

Before going on, I should declare my own philosophical stance. My perspectives on e-learning have been influenced by 25 years of explorations into applying artificial intelligence to education (AIED). The AIED approach is fundamentally focused on two things: the need for e-learning systems to be adaptive to individual differences in learners, and the belief that advanced computational technologies (in particular artificial intelligence) need to be adapted and/or developed to achieve this goal. The overarching goal of most of my work has been to support learners as they learn, for example doing their assignments, working on projects, or carrying on some authentic activity in the workplace. In collaboration with colleagues and graduate students, innovative educational technologies have been built and deployed (with greater or lesser degrees of success) with real learners. While using (and developing) advanced computational techniques, this work has also drawn inspiration and ideas from areas such as situated learning, collaborative learning, instructional design, constructivist learning, virtual learning communities, etc. I summarize my perspectives on how all of this ties together in an issues-oriented paper (McCalla 2000) outlining what I believe to be important future directions for research into e-learning systems.

In this paper I will make arguments for an approach to the design of e-learning systems, called the *ecological approach*, that shows promise to allow information about how learners use a system to be naturally captured and then used in a multitude of ways. In a phrase, the approach involves attaching models of users to the information they interact with, and then mining these models for patterns that are useful for various purposes. The information and the data mining algorithms interact with one another in an ecosystem where the relevance and usefulness of

information is always being adjusted to suit the changing needs of learners and teachers and to fit changes in the external environment and the system's perceptions. I will draw on ideas and techniques from AIED, user modelling, collaborative filtering, case-based reasoning, the semantic web, data mining, multi-agent systems, information retrieval, recommender systems, the learning objects movement, cognitive science, instructional design, and other social sciences. Particularly important are learning objects in which learning activities and material are encapsulated; the semantic web with its notions of user-centric open access and metadata to expand usability; learner modelling with its focus on individual learners and adaptivity to their needs; collaborative filtering for its concentration on similarities among users; and data mining to make sense of large amounts of unstructured data.

The paper is organized as follows. In section 2, I will motivate and outline the ecological approach to capturing and using information about users that shows great promise particularly in e-learning contexts. In section 3, I discuss current research projects with which I am involved that are exploring several of the issues raised by the ecological approach. In section 4, I look at some related research disciplines and discuss the possible implications of the approach advocated here on these disciplines. Finally, in section 5 I summarize the advantages of the ecological approach and the hurdles that must be overcome if the approach is to fulfill its potential.

2.   The Ecological Approach

2.1   Overview of the Approach

In most semantic web research and development it is assumed that web content is marked up using standardized metadata. Such metadata is meant to add information that search engines and other Internet technologies can use to more accurately "understand" the pages and retrieve and/or manipulate them for the user. There has been a considerable international effort to develop metadata standards and tools that allow computation and interoperability on information that adheres to these standards. In the education domain alone there are standards such as the Dublin Core Metadata Initiative (DCMI 2003), ARIADNE (2000), IMS (Anderson and McKell, 2003), IEEE LOM (IEEE 2002); domain and culture specific schema (CanCore (Fisher, *et al.* 2002), Le@rning Federation Profile (Friesen, *et al.* 2002); educational modelling languages (such as EML (Koper 2000)); and tools (such as the POOL-POND-SPLASH suite (Hatala and Richards, 2003). Once such standards are

agreed upon, then it is assumed that content providers will mark up "content" using the standard metadata tags and that anybody wishing to access the content will use the same tags, thus providing additional semantics to the content.

Unfortunately, there are a number of problems with this approach. In particular, there is no way of guaranteeing that the metadata really capture the various domains with the breadth and depth needed. Moreover, the metadata tend to be about the form and the content of the page, even though many other kinds of information could be useful for various purposes. A related problem is that the same page can be used for several purposes, and that the metadata relevant for one purpose may be different from that required for another purpose. There is also no way to guarantee consistency in the application of metadata to content: different content developers, different users, and different applications may interpret the same metadata in different ways. There is also a heavy front-end load in the standard approach, with the requirement for pre-assigning metadata to content to make it usable.

When the domain is education, there are a number of special problems with the standard approach. It is important in educational applications to understand the individual needs of each learner, and yet content or form-based metadata have little role for distinguishing one user from another. Ideally, learning objects need to reflect appropriateness to differences among learners' cognitive development, learning styles, motivation and other affective characteristics, etc., in addition to content. They also need to incorporate aspects that allow pedagogical decisions to be made, including information about prerequisites, level of detail, technical level, etc. The standard approach also doesn't allow very well for change. Not only does the content change, learners, by definition, are constantly changing as they gain mastery: their use of a given learning object may differ substantially depending on their stage of learning, the same object fulfilling different roles at different stages. Also, an e-learning system's understanding of learners is constantly changing with more interactions between learners and the system.

To overcome these problems, my colleagues, graduate students, and I are working on an alternative to the standard approach called the *ecological approach*, essentially an enhancement of collaborative filtering approaches. In the ecological approach, information about web content is attached to the content as users access that content. The information may include:

- information about the users, including cognitive, affective, and social characteristics of the users and their goals in accessing the content;
- information about the content itself, including the users' opinions of what

the content is and interpretations of the content inferred by text processing algorithms (eg. through latent semantic analysis), and pre-specified metadata from a known ontology;

- information about how the users interacted with the content, including observed metrics such as dwell time, number of user keystrokes, patterns of access, etc., and users' opinions of the effectiveness of the content in meeting their goals;
- information about the technical context of use, including characteristics of the users' software and hardware environment;
- information about the social context of use, including access to a particular user's previous experiences with other content and access to other users' experiences with this content.

As time goes on, more and more users will access the content and more and more information will thus accumulate about the content. For the most part, this information will be about real users' real use of the content, gathered real time during their use. It will not be information from a standard ontology added by some external person prior to the use of the content. This gradual accumulation of information and the focus on end use are two key aspects of the ecological approach.

The third key aspect of the ecological approach is the *purpose-based* use (in the sense of McCalla, Vassileva, Greer and Bull, 2000) of the information associated with the content, to achieve a particular goal. There are many purposes that would be better fulfilled with a deep and broad understanding of content, including recommending relevant content for a particular user with a particular need for it, tailoring the content to a particular user's goals and/or needs, evaluating the effectiveness of the content in meeting the needs of various types of users, deciding whether the content is still relevant or has become obsolete, determining semantic relationships (eg. similarity, abstraction, aggregation, etc.) between the content and other content, etc. Each such purpose places its own particular constraints on what information is relevant and how it is to be used to help to achieve the purpose. Thus, determining whether to recommend specific content to a particular user may require comparing this user to other users on important characteristics and then looking at how similar users have evaluated the content (and, moreover, the characteristics considered to be important are themselves determined by the user's goals). On the other hand, determining whether the content is now obsolete may require an examination of all users' evaluations of the content, trying to extract temporal patterns in the evaluations that show how recent users like or dislike the content. Important technologies supporting this kind of purpose-based use of information are data

mining and clustering techniques from artificial intelligence. The key point is that it is the purpose that determines what information to use and how it is to be used. Further this determination is made *actively* (in the sense of McCalla, Vassileva, Greer and Bull, 2000) at the time the purpose is invoked: no a *priori* interpretation needs to be given to the information.

In sum, then the ecological approach promotes the notion that information gradually accumulates about content, the information is about the use of the content by real users, and this information is interpreted only in the context of end use. The approach is ecological because over time the system is populated with more and more information, and something like natural selection based on purposes determines what information is useful and what is not.

2.2   The Ecological Approach in E-Learning

For some applications, it will be difficult or impossible to have all the relevant knowledge of users or their goals, thus limiting the effectiveness of the ecological approach. However, in educational domains virtually all of this knowledge can be readily available. Traditionally, learners have proven to be more willing to provide information to systems that will help them to learn than they have been to standard application systems, aimed at commercial profit for others, especially if they think that such information will make the e-learning system more effective and responsive to their own needs. They are also more likely to be willing to be monitored and evaluated, including to allow diagnosis of their problem solving behaviour (to facilitate intelligent feedback) and the testing of their knowledge. Educational goals can be explicitly known; for example a learner may want to learn about some subject, to find content relevant to a particular issue, to get help to overcome an impasse, etc. Thus, the educational domain is an excellent place to explore the ecological approach, since there is every chance of a very high bandwidth of interaction between the learner and the system.

In the educational domain there are many possible applications for the ecological approach. The approach could underlie the design of:
- a study aid, for example to retrieve for a learner relevant papers from a cache of such papers for a graduate student trying to learn about an area of research (as in Tang and McCalla, 2003);
- a recommender system, to recommend some content to a learner that is relevant to his or her current task (as in Recker and Wiley, 2001);
- an instructional planner, to plan out a sequence of content pages of relevance

to a learner, sort of an individualized curriculum of study;
- a group formation tool, to suggest to the learner a group of other learners relevant to solving a particular task or learning about a particular subject;
- a help seeker, to find another learner who can help the learner solve a problem he or she has encountered (as in I-Help, Greer, McCalla, Cooke, Collins, Kumar, Bishop and Vassileva, 1998);
- a reminder system, to keep a learner updated with new relevant information, say from the web, that is relevant to the learner's goals;
- an evaluation tool, to allow learners' interactions with educational content to be studied by instructional and cognitive scientists, in particular to look at the experiences of all learner or particular types of learners with some educational content;
- an end-use tagging system, to automatically derive educational content tags from pre-established ontologies based on the experiences of the actual users of the content, and that can be parameterized by end use variables such as type of learner, success/failure of the educational content for each type of learner, etc.
- an "intelligent" garbage collection system, to determine the on-going relevance of educational content and, if necessary, to suggest modifications or even that it be deleted as no longer being useful to learners (eg. as discussed in Bannan-Ritland, Dabbagh and Murphy, 2000).

In each of these potential applications it is assumed that the system has access to a *repository* of educational content (eg. the MERLOT (2003) or CAREO (2003) repositories) annotated with characteristics and experiences of the various learners who have accessed the content. The content could be web pages, research papers, courses of study, text books, etc. The term *learning object* will stand for any of these possibilities. It will also be convenient to consider that the learner characteristics and experiences are packaged as *learner models* for each learner, in the AIED sense of the term. Thus, the ecological approach for educational domains boils down to this: each time a learner is interacting with a learning object, the learner model for that learner (in its current instantiation, of course) is attached to the learning object. Over time this means that each object collects many learner models and these can be mined for patterns of particular use to a given application.

The semantics-based paradigm of externally tagging content with pre-established labels from a standard metadata set is thus transformed to a *pragmatics*-based paradigm of tagging content with learner models which can be reasoned with in the context of use. Such a transformed focus allows the end purposes of the application

and the particulars of the learning context (including the specific learner(s) involved) to drive the inferences made by the application, with a consequent increase in the power and sensitivity of the application. Perhaps we should be talking about the pragmatic web, not the semantic web!

The key technologies underlying the ecological approach are user modelling, in particular for educational domains learner modelling, and data clustering and data mining. First, consider the learner modelling issues. The learner model can be roughly divided into two parts. In one part, the *characteristics* part, will be information about the learner that transcends any given activity of the learner. This would include personal information like name, gender, age, etc.; affective and social information; preferred cognitive style and/or learning style; learning goals/objectives for this learning episode; previous learning objects that the learner has interacted with; indicators of success in learning such as marks, awards, etc. In the other part, the *episodic* part, will be information about the learner's experiences with the current learning object. This would include traces of the learner's interactions with the learning object (possibly down to the keystroke level); the learner's evaluation of the learning object on various criteria such as level of difficulty, technical level, etc.; the learner's opinion of primary and secondary content (expressed, perhaps, in standard ontologies for the domain); results of tests of competence on the content of the learning object (note that in educational domains such evaluation is more natural and easier to motivate for the learner than in most other domains such as e-commerce). Information in both parts can change over time, although obviously the information in the episodic part is completely different for each learning object visited, so must be reconstructed from scratch whenever the learner starts interacting with a new learning object (but the old episodic information is not thrown away, as will become evident next).

When a learner is accessing a learning object, both parts of the learner model can be used in standard fashion to inform the educational interaction between the learner and the learning object. Once the learner has finished, the learner model in its current state is copied and attached to the learning object as a *learner model instance*. Of course, this learner model instance does not change once the learner has moved on: even as the learner model itself evolves and changes with the learner as the learner interacts with other learning objects, the learner model instance stays behind as a snapshot of the learner's experiences with this particular learning object at this particular time and in this particular context. When other learners interact with this learning object (or the same learner returns in the future), their learner model instances are also attached. Incrementally, more and more learner model instances

accumulate which should allow more and more refined reasoning about the learning object's actual implications for learners. In addition to learning object instances, also attached to a learning object, of course, could be standard metadata assigned by pedagogues or other professional indexers from standard ontologies (eg. IMS-LD or the IEEE-LOM standards). It is an interesting possibility of the ecological approach to allow this pre-assigned metadata to be refined, modified, or even changed based on inferences from end use; that is pre-assigned metadata about content may not agree with what the learners thought the content was, or at least not for every type of learner.

The other technologies of importance are data clustering and data mining. Once a sufficient number of learner models has been attached to the various learning objects in a learning object repository, it is possible to use these technologies to find interesting and relevant patterns in the information contained in the learning object instances. What is interesting and relevant, however, is not absolute, but is instead relative to the educational application and the particular purpose that the educational application is trying to achieve. Consider, for example, a recommender system trying to choose a relevant learning object from a learning object repository that would be appropriate for a particular learner, L. First, the recommender system would access L's current learning goal(s) (from the characteristics part of L's learner model, L_m). Such goals would indicate things such as the content the learner is interested in understanding and the desired depth of understanding. The recommender system then would retrieve (also from the characteristics part of L_m) information relevant to these learning goal(s) such as L's learning style, affective and social aspects of L, and the previous learning objects L has visited. Using an appropriate clustering algorithm, this information would then be compared with the same characteristics within learner model instances attached to the learning objects in the repository, to retrieve a set of *similar* learner model instances {L1, L2, …, Ln}. Next the episodic part of each of the $L_i \epsilon \{L_1, L_2, …, L_n\}$ would be consulted to gather data about the experiences of the learner associated with each L_i, such as results of tests, activity measures like dwell time, learner evaluations, etc. For each L_i, this experiential information would be summarized to determine how well the learning object had served its learner. When the summarized experiential information is highly positive, the associated learning object probably should be recommended to L since a similar learner had a positive and useful experience with the learning object. But, when the experiences of similar learners is negative, it is unlikely that the learning object will be helpful to L.

Consider another example application: an evaluation system meant to determine the strengths and weaknesses of a learning object for achieving a particular pedagogical objective. This evaluation system would first retrieve all of the learner model instances $\{L_1, L_2, \ldots, L_m\}$ attached to the learning object. An appropriate data mining algorithm would then be fired up to look for patterns connecting these learner model instances. The data mining algorithm would focus its search for information and patterns relevant to the pedagogical objective. Patterns associated with different types of learners (determined by having similar characteristics) could allow evaluations of the effectiveness on different groups of learners, as well as a summary of overall strengths and weaknesses in achieving the pedagogical objective.

The key to both of these examples (and this would be true for all of the proposed applications listed above for the ecological approach) is that they follow the active and purpose-based learner modelling approach promoted by McCalla, *et al.* (2000) and Vassileva, *et al.* (2003). That is, contextual information such as the purpose, the particular learner(s) involved, the application goals, etc., determine how (and even whether) the information in the learner models and learner model instances is used. The choice of clustering algorithm and/or data mining algorithm, and the particular constraints put on each such algorithm is highly contextualized. Many other algorithms are also similarly contextualized, such as the experiential summarization algorithms in the recommender system example. Thus, much research is needed into what algorithms work, where, and for what purposes. As in all active approaches, there are many space/time tradeoffs that must be resolved. How much pre-computation can be done to find patterns that can then be retrieved quickly when real time response is needed? How can such pre-computation be done before the various contextual elements are known? How much information can be kept around in learner model instances before there is too much information to deal with? Can this information be compressed or deleted or summarized while still allowing finely tuned performance? There are other serious problems to be resolved too. How can the ecological approach still work in the early stages before there has been much learner interaction with the learning objects in a repository (this is a version of the cold start problem faced by many case-based systems)? Obviously, there is a vigorous research agenda that lies ahead!

There is also much research needed into the structure of the learner models, the kinds of information that can be gathered, the kinds of information that are useful (and for what purposes), etc. After some experience in applying the ecological approach in a wide variety of applications and a diversity of situations, it will perhaps be possible to devise standard learner model slots and slot fillers, and standard ways

of using the information in these slots to achieve particular purposes. This would move the standardization efforts away from defining vocabularies of terms with which to index learning objects, to defining standards for learner models to be attached to learning objects and standards for ecological inference procedures. This would allow interoperability and reuse to be achieved even if an external learning object repository were imported. An additional approach to interoperability and re-use is also being explored in our laboratory in the LORNET project, discussed below.

3. Specific Projects Investigating Issues of Importance to the Ecological Approach

Together with my research colleagues and our graduate students I am investigating a number of issues related to the ecological approach. The foundation of the ecological perspective is our experience with the I-Help system (Greer, *et al*. 1998). I-Help is an open peer system to support learning, developed over the last decade in our laboratory by a number of faculty (in particular Jim Greer, Julita Vassileva, Ralph Deters, John Cooke, and myself), research associates, and graduate students. I-Help is focused on helping learners as they learn, and on deploying learners themselves in both helping and being helped mode. It has been deployed in computer science courses at the University of Saskatchewan and in other contexts to many thousands of students over the last 5 years (McCalla, Greer, Vassileva, Deters, Bull, Kettel, 2001).

I-Help has two components: the widely used public discussion (I-Help Pub) component and the less well used, but more interesting, private discussion (I-Help 1-on-1) component. In I-Help Pub, learners can post questions, comments and responses to forums. These postings are shared with their peers. Forums are clustered into groups and group memberships. A person who is a member of a group can access the forums created for that group. I-Help Pub is used asynchronously.

The second I-Help component, I-Help 1-on-1, supports one-on-one private discussions (or help dialogues) between a learner requesting help, the helpee, and a single peer (or expert), the helper. These dialogues may be synchronous or asynchronous. In I-Help 1-on-1, a learner contacts their personal agent to issue a help request. The learner's agent negotiates with the agents of other learners, to locate potential helpers. The top N matches are notified that there is a help request waiting. The first of the contacted helpers to accept the request starts a one-on-one interaction with the helpee. Requests to other potential helpers are cancelled. Upon completion of the interaction, each learner receives a brief evaluation form through

which they evaluate their partner, and this can be added to the personal agents of both the helper and the helpee.

Multiple fragmented learner models underlie the I-Help 1-on-1 system. Each person is represented by a personal agent in the system, and this personal agent keeps a model of its "owner" (and possibly other learners) as a source of information as it acts on the owner's behalf. These models are used by personal agents when negotiating help sessions with other users in order to determine the best helpee-helper matches. Learner model information is obtained in a variety of ways: from the learner (through stated availability and self-assessment of knowledge of different topics); from the short peer evaluations; from a determination of whether or not the learner is currently or frequently online; and from I-Help's observations of learner partici-pation in both the public and private discussions. The public and private discussions may be used together, or the two components may be used independently. Whichever is used, the obvious educational benefit to learners is that those requiring help receive assistance at the time they need it. Furthermore, peers providing help should also benefit from the reflection necessary to formulate an acceptable explanation.

As an outgrowth of our work on I-Help, we have begun to define and explore a new learner modelling paradigm we have called *active learner modelling* (McCalla, *et al.* 2000; Vassileva, et al. 2003) because the focus is on issues to do with computing partial learner models in context, rather than with the representation of a compre-hensive single learner model. This paradigm fits the agent-based distributed computational environments that are increasingly prevalent in information technology, including educational environments such as Edutella (Nedjl, *et al.* 2002) or COMTELLA (Bretzke and Vassileva (2003). In the active paradigm, raw information about learners is slowly gathered over time by a number of agents (most importantly personal agents representing learners). The information about any given learner is thus fragmented among many agents. This raw information is then actively interpreted by particular application agents to achieve some purpose. The purpose defines which information is relevant, where to look for it, how to combine it, and what sense to make of it. Not coincidentally the active learner modelling perspective meshes well with the ecological approach. In one experiment (carried out by Julita Vassileva's and my M.Sc. student Xiaolin Liu – see Niu, McCalla and Vassileva, 2003) we have focused on defining purpose hierarchies in particular domains and designing modelling algorithms for each purpose. These algorithms are anytime algorithms, in that computation can be stopped at any point as resources and time permits and there will still be some model computed. The more time and resources

there are, the more refined will be the model, of course, but in many real world applications, with real time constraints, these will be in limited supply. This investigation has been carried out in a simulated stock investment domain, but the lessons transfer readily to e-learning environments.In another experiment in active modelling, we (Jim Greer, me, and our summer student Ryan Silk) have developed a set of data mining algorithms to find interesting patterns in the I-Help public discussion forums to further inform personal agents about characteristics of potential helpers for a learner needing help. Finally, Susan Bull and I have been investigating cognitive style and how to capture and use it in the I-Help context (Bull and McCalla 2002). Extending the range of learner characteristics that can be modeled can add new dimensions to the patterns discovered by the data mining algorithms.

In a project with direct overlap with the ecological approach, my Ph.D. student, Tiffany Tang, is building an advising system for graduate students who are trying to read the literature in some potential area of research (Tang and McCalla, 2003). In this situation, learning objects are research papers, and the learning object repository is a set of current papers in a particular research area. Her system has three basic purposes: (i) to recommend a paper to a particular graduate student that is relevant to his or her current needs; (ii) to allow the student to annotate his or her impressions of the paper once it has been read; and (iii) to keep the paper repository up to date as new papers appear and old papers become irrelevant. Her approach is directly ecological; in fact, this is the research project that defined many of the specifics of the ecological approach. A model is kept of each graduate student, and an instance of that model is attached to each paper after the graduate student has read the paper. As in the above description of the ecological approach, the learner model instance has two parts: the student characteristics part, directly inherited from the model; and the episodic part, derived from student annotations providing the student's impressions of the paper. Purpose (i) is achieved by finding a cluster of student model instances having similar characteristics and episodic behaviour to the current student, and recommending papers that this cluster of students likes (using a clustering algorithm devised by Tang, Chan, Winoto and Wu, 2001). Purpose (ii) is achieved by providing a tool that graduate students will find useful for annotating a paper, one that allows them to indicate things like the paper's content, technical level, comprehensibility, depth, and breadth (by selecting from a pre-established menu of terms); to write a brief summary of the paper; and to retrieve the full paper citation. Note that this purpose not only serves the student but also serves the system by feeding episodic information into the learner model instance. Retrieving new papers (one part of purpose (iii)) is achieved by searching CITESEER. Deleting old papers (the other part of purpose (iii)) is achieved by looking at recent student model

instances for papers that no longer are being evaluated positively. Purpose (iii) is not a central focus of the current investigations in this project (purpose (i) is the main current focus), but it has many interesting research implications. More sophisticated and intelligent examination of the papers in the repository through the lens of the students' experiences with them could allow all manner of inferences to be drawn about papers, including what they are about, how they relate to each other, how the research discipline is changing, what papers appeal to what types of readers, etc. This project is now being precisely specified and a prototype "proof of concept" implementation should be ready within about 6 months from this writing.

A final project of relevance to the ecological approach is a recently launched Canadian national collaborative investigation of learning object repositories called LORNET. Directed by Gilbert Paquette of the TeleUniversité du Québec, and with researchers at the University of Ottawa, University of Waterloo, Simon Fraser University, and the University of Saskatchewan, many different aspects of learning object repositories are being explored. The principal investigators for the University of Saskatchewan LORNET group are Gord McCalla, Jim Greer, Julita Vassileva, Ralph Deters, and John Cooke. Our goals in LORNET are to turn learning objects into active agents, able to negotiate their interactions with other learning object agents, personal agents representing learners, and computational processes. This is essential, we feel, for learning objects to be re-used in new contexts for a specific learner or learners. The subtle interactions and adjustments necessary for collections of learning objects to serve the needs of a learner or group of learners can best be handled by the learning objects and personal agents for the learner(s) interacting with one another. Our focus is on end use, and how learning objects can be personalized to the individual needs of the learners. The learner modelling approach is active, in the sense that the focus is on computing models for particular purposes, and the systems we are building are ecological in the sense that learner modelling information accumulates with each learning object and can be used by the various agents in their negotiations and their other activities. We are optimistic that our part of LORNET will shed light on fairly deep issues of learning object personalization, adaptation, and re-use. It will also be another exploration of the ecological approach.

4.   Implications of the Ecological Approach for Research

The ecological approach is sympathetic with a number of directions being explored in the educational and technological research communities:

- educational
    - o   recent approaches to instructional design
    - o   the learning objects movement

- technological
    - o   the semantic web
    - o   various computational techniques, most notably collaborative filtering and data mining

In this section I would like to discuss how the ecological approach is impacted by and impacts these various research communities.

The ecological approach most naturally supports learner-centered constructivist pedagogical philosophies (as described, for example, in Bannan-Ritland, *et al.* 2000), although any philosophy could be supported. The focus in the ecological approach is on learners engaged in authentic learning activities being supported by technology to achieve their goals.

Personalization and individualization are desirable, but the approach also supports collaboration and interaction (as the I-Help and LORNET projects illustrate). A project with closely related goals and philosophy is that of Recker and her colleagues in the Reusability, Collaboration, and Learning Troupe (RCLT) at Utah State University, who place explicit emphasis on constructivist and community-based learning paradigms, peer-to-peer tools, and collaborative filtering to factor in user experience (Recker and Wiley, 2001). The collaborative filtering in the ecological approach is a generalization of that in Recker and Wiley and other collaborative filtering approaches, in that it posits full scale evolving learner models to capture a broad range of learner characteristics and end use experience. Also, in the ecological approach the explicit representation of purposes allows purpose-specific data mining and clustering algorithms to carry out appropriate computations on data relevant to that purpose. This provides a high degree of flexibility to the approach, and allows a wide variety of purposes, not just recommending useful learning material, to be carried out.

An interesting instructional design research implication of the ecological approach would be to map the pedagogical principles of a particular education philosophy onto purposes carried out by computations that would in some sense enforce the paradigm. Such computations would draw conclusions and provide support appropriate to the particular pedagogical paradigm, but would still be learner-centered since the computations are directly operating on a knowledge base of learner characteristics and experience. In this sense (and appropriately) any pedagogical paradigm would be learner centered, even instructivist paradigms. Work such as that of Martinez (2000), which studies how various factors about the learner (including learning style and affective and social characteristics) can affect the choice of appropriate pedagogy, could form an intermediate level of analysis that would be very useful in mapping from pedagogical paradigm to the computations needed to implement that paradigm in an e-learning system.

Another major educational influence on this research is the learning objects movement. The ecological approach draws much from investigations into learning objects, including encapsulating learning resources into objects adorned with metadata, collecting learning objects into learning object repositories, and supporting re-use and interoperability. However, the ecological approach makes several transformations to the standard learning object paradigm. First, the metadata are learner models, rather than terms from standard ontologies, and the metadata are added automatically as learners interact with the learning objects. This allows the capture of end use data without the need for a human to pre-attach metadata. Second, this metadata is not given any a priori significance, but is instead actively interpreted in the context of the particular purpose and the particular learner(s) involved. This means that the metadata can mean different things depending on the context. The same learning object could be "about" entirely different things, and have entirely different pedagogical implications, for learners with different goals. Meaning is context-dependent and relative to the purpose at hand. Appropriate adaptation to context is a crucial element in making learning objects re-usable (Wiley, 2003). As Anderson and Mah (2002) note in regard to a project where they extensively experimented with the reuse of learning objects: "…trying to decouple knowledge from context, and then implement them in context-driven needs was a futile task."

These differences have interesting implications for learning objects research. More effort should be redirected towards finding standards to describe learner models (standard kinds of learner characteristics, standard types of learner activity – a start has been made in the PAPI standard (IEEE, 2000)), and to define a range of standard

purposes and the accompanying computations to carry them out. The ecological approach also mandates the development of a different array of computational tools, including tools to clean up the learning object repositories of irrelevant or ineffective learning objects, tools to summarize learner behaviour, tools to abstract common-alities among learner models, procedures to mine and cluster learner model data, tools to prune the information in learner models to a manageable level, etc. The focus of the ecological approach on end use context suggests a different way of supporting learning object re-use, a problem for the standard approach to learning objects (RCLT, 2001). Instead of de-contextualizing an object so it can be used in any context, the ecological approach effectively suggests keeping context information with the object and then matching any new context to previous contexts of use to determine whether (and to some degree how) to re-use the learning object. Another aspect of re-use is allowing learning objects to negotiate their interrelationships, essentially making them into learning agents (as we are exploring in the LORNET project). By having learner models of the actual learners attached to the learning objects (agents) to provide the agents with a wide range of end use information to work with, by having a particular learner or learners represented by personal agents also actively involved in the negotiations, and by having an overriding pedagogical purpose for the negotiations, groups of learning objects (agents) have the potential to adjust themselves specifically to a given situation and set of learners. Of course, there are a very large number of issues to be resolved in actually making this learning agent approach work.

A major technological area that has influenced the ecological approach is obviously the semantic web. The discussion about learning objects largely applies also to differences between the ecological approach and standard approaches to the semantic web: metadata that are user model instances not terms from standard ontologies, the slow accretion over time of user model instances rather than the need for pre-tagging by a human expert, actively making sense of the metatdata only in the context of a particular purpose and for particular user. But can the ecological approach generalize from learning situations to the open web? Learning environments are well suited for the ecological approach: the environment can be constrained to a limited set of learning objects; each learner is engaged over a considerable period of time and so more can be known about him or her as time goes on; it is possible to know many characteristics of learners; and learners are likely to be more willing to state their goal (or be provided with one), to be monitored, to take tests, and to undertake activities under direction and receive advice for the greater good of learning a subject. It is unclear whether the fully open web, where it is hard to know much about users, their goals are hugely varied and always changing, and where they are not usually willing

to provide much direct feedback to the system, is well suited to the ecological approach. It is surely a very much more complex situation, but the ecological approach, in principle at least, might still work if enough bandwidth of interaction with users can be gathered and maintained over a long enough time. There has been work on adaptive hypermedia (Brusilovsky, Kobsa and Vassileva, 1998) and person-alization on the web (Vassileva, McCalla and Greer, 2003; Dolong and Nejdi, 2003) that could provide insights into the possibility of tackling the hard problem of applying the ecological approach to the open web.

There are many technologies that are needed for the ecological approach to be successful. Most central are data clustering, data mining (in particular usage mining, Pierrakos, Paliouras, Papatheodorou, Spyropoulos, 2003), collaborative filtering, and learner modelling (in particular active learner modelling). Apart from learner modelling, of course, few of these techniques have been used in e-learning, but this seems to be changing now: for example, Zaizne and Luo (2000) have used data mining in an e-learning context, to mine learners' activities and the pages they have browsed in order to recommend new pages to them. There is much empirical research ahead to determine which of these algorithms can be useful where and when in implementing the ecological approach, and perhaps even the need to develop new algorithms, many of which will have to have an anytime flavour given the usual need for real time response to learners. There are also big computer science problems: finding efficient mining and clustering algorithms; making space-time tradeoffs, in particular determining what patterns can be found offline for easy real time retrieval later (and, on the other hand, what patterns can only be looked for during interaction with the learner); deploying learning algorithms and other techniques that allow information to be generalized and abstracted from the plethora of learner models that will increasingly clutter the system if not pruned; and intelligent garbage collection of learning objects when they are no longer useful or relevant.

5.   Conclusions

This paper has argued that e-learning systems could be ecological in the sense that they could continuously be adapting as the e-learning system's understanding of its external environment changes and as the external environment itself changes. The external environment includes learners, teachers, the subject matter being learned, and the technology that implements the e-learning system. The adaptation includes the possibility of modifications to the objects in the e-learning system, the possible deletion of some objects, and/or the addition of new objects. Over time, then, the e-learning system slowly evolves, fine tuning itself to its environment and keeping abreast of change in that environment.

The ecological approach proposed here has a number of explicit features: it focuses on end-use; it provides a natural way of capturing end-use information; it has a central role for learners and their goals; it is contextual, where context is most importantly a function of purpose and the people (learners and teachers) involved in the learning situation; it is procedural in its emphasis on the process of making sense of information in context; it has need for knowledge of individual learners but uses this to support all learners and to make system level decisions; it naturally supports constructivist learner-centered pedagogical principles; it allows an e-learning system to incrementally evolve and adapt as its environment changes and as it knows more about the environment. The approach scales well as the number of learners grows; in fact, the more learners, the *better* the ecological approach is likely to perform. The ecological approach draws inspiration from many research communities, including various e-learning paradigms (especially artificial intelligence in education and learning objects), user modelling, the semantic web, collaborative filtering, data mining, and instructional design. It also suggests a number of new research directions including the study of purposes; the exploration of data mining and clustering algorithms that can find patterns in learners' behaviour; efforts aimed at standardizing these algorithms and standardizing the learner model structure; the exploration of a specific notion of context based on purposes and the people involved in the learning situation; the study of intelligent garbage collection; and deep investigations of computational issues such as computational complexity and space-time trade-offs. Many of these issues are very hard, and it is unclear how widely the ecological approach can be applied, even though the approach seems very promising.

The final take-away lesson of this paper is that a valuable direction for research is to look at end-use context and thereby to transform investigations of the semantic web into investigations of the pragmatic web.

Acknowledgements

6.   References

Anderson, A. and Mah, S. (2002). Keeping It Relevant: An Object Approach to Training Package Content Development That Facilitates Workplace Contextualisation. Annual Conference of the Higher Education Research and Development Society of Australasia, July.

Anderson, T. and McKell, M. (2003). IMS Learning Resource Meta-data Information Model, Version 1.2. IMS Global Learning Inc.
<http://www.imsglobal.org/metadata/imsmdv1p2/imsmd_infov1p2.html>

ARIADNE (2000). The ARIADNE Foundation for the European Knowledge Pool <http://www.ariadne-eu.org/>

Bannan-Ritland, B., Dabbagh, N. and Murphy, K. (2000). Learning Object Systems as Constructivist Learning Environments: Related Assumptions, Theories and Applications. In D. Wiley (Ed.), *The Instructional Use of Learning Objects (on line version)*, section 2.1, Association for Instructional Technology and Association for Educational Communications and Technology (AIT/AECT).
<http://www.reusability.org/read/>

Bretzke H. and Vassileva J. (2003). Motivating Cooperation in Peer to Peer Networks. Proceedings UM 03: International Conference on User Modeling, Johnstown, PA, June Springer-Verlag: Berlin.

Brusilovsky, P., Kobsa, A. and Vassileva, J. (eds.) (1998). *Adaptive Hypertext and Hypermedia*. Kluwer Academic Publishers Dordrecht.

Bull, S. and McCalla, G. (2002). Modelling Cognitive Style in a Peer Help Network. Instructional Science Journal, 30(6), 497-528.

CAREO (2003). Campus Alberta Repository of Educational Objects.
<http://www.careo.org/>

DCMI (2003). Dublin Core Metadata Initiative. <http://dublincore.org/>

Dolong, Peter and Nejdi, Wolfgang (2003). Challenges and Benefits of the Semantic Web for User Modelling. Learning Lab, University of Hannover, Germany.

Fisher, S. *et al*. (2002). CanCore Learning Object Metadata: Metadata Guidelines version 1.1, CanCore Initiative, Athabasca University, Edmonton, Alberta, Canada.

Friesen, N. *et al*. (2002). Building Educational Metadata Application Profiles. Proceedings of International Conference on Dublin Core and Metadata for e Communities, 63-69, Firenze University Press, Italy.

Greer, J., McCalla, G., Cooke, J., Collins, J., Kumar, V., Bishop, A., and Vassileva, J. (1998). The Intelligent Helpdesk: Supporting Peer-Help in a University Course, Proceedings of ITS 1998: Fourth International Conference on Intelligent Tutoring Systems, 494-503, August, San Antonio, Texas, Springer-Verlag: Berlin, Germany.

Hatala, M. and Richards, G. (2003). Making a SPLASH: A Heteregeneous Peer-to-Peer Learning Object Repository. Proceedings of WWW 2003: The Twelfth International World Wide Web Conference, Association of Computing Machinery (ACM).

IEEE (2000). IEEE P1484.2/D7, 2000-11-28, Draft Standard for Learning Technology, Public and Private Information (PAPI) for Learners (PAPI Learner). <http://ltsc.ieee.org/wg2/>

IEEE (2002). IEEE P1484.12.1-2002, Draft Standard for Learning Object Metadata, IEEE, Inc.

Koper, R. (2000). From Change to Renewal: Educational Technology Foundations of Electronic Environments. Manuscript of Speech, White Paper, Open Universiteit Nederland.

Martinez, M. (2000). Designing Learning Objects to Mass Customize and Personalize Learning. In D. Wiley (Ed.), *The Instructional Use of Learning Objects (on line version)*, section 3.1, Association for Instructional Technology and Association for Educational Communications and Technology (AIT/AECT). <http://www.reusability.org/read/>.

McCalla, G. (2000). The Fragmentation of Culture, Learning, Teaching and Technology: Implications for the Artificial Intelligence in Education Research Agenda in 2010. International Journal of Artificial Intelligence in Education, 11(2), 177-196.

McCalla, G., Greer, J., Vassileva, J., Deters, R., Bull, S., Kettel, L. (2001). Lessons Learned in Deploying a Multi-Agent Learning Support System: The I-Help Experience, Proceedings of AIED 2001: International Conference on AI in Education, 410-421, May, San Antonio, Texas, published as J. Moore, C. Redfield, and W.L. Johnson (eds.), Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future, IOS Press: Amsterdam.

McCalla, G., Vassileva, J., Greer, J. and Bull, S. (2000). Active Learner Modeling. Proceedings of ITS 2000: Intelligent Tutoring Systems, 53-62, June 2000, Montréal, QC, Canada. Springer-Verlag: Berlin, Germany.

MERLOT (2003). Multimedia Educational Resource for Learning and On-Line Teaching. <http://www.merlot.org/>
Nejdl, W. et al. (2002). EDUTELLA: A P2P Networking Infrastructure Based on RDF," Proceedings of WWW 2002: The Eleventh ional World Wide Web Conference, Association of Computing Machinery (ACM), Honolulu, Hawaii.

Niu, X., McCalla, G. and Vassileva, J. (2003). Purpose-based User Modelling in a Multi-Agent Portfolio Management System. Proceedings of UM 2003: International Conference on User Modeling, 398-402, Johnstown, Pennsylvania.

Pierrakos, D., Paliouras, G., Papatheodorou, C., Spyropoulos, C. (2003). Web Usage Mining as a Tool for Personalization: A Survey. J. User Modeling and User Adapted Interaction, 13, 311-372, Kluwer Academic Publishers: Netherlands.

Recker, M. and Wiley, D. (2001). A Non-Authoritative Educational Metadata Ontology for Filtering and Recommending Learning Objects. Interactive Learning Environments Journal: Special Issue on Metadata, 1-17.

RCLT (2001). The Re-Usability Paradox. White Paper, The Reusability, Collaboration, and Learning Troupe, Utah State University, Utah.
<http://rclt.usu.edu/whitepapers/paradox.html>

Tang, T., Chan, K., Winoto, P. and Wu, A. (2001). Forming Student Clusters Based on Their Browsing Behaviors. Proceedings of ICCE 2001: 9th International Conference on Computers in Education, 1229-1235, November 2001, Seoul, Korea.

Tang, T. and McCalla, G. (2003). Smart Recommendation for an Evolving E-Learning System. Proc. Workshop on Technologies for Electronic Documents for

Supporting Learning, Int. Conf. on Artificial Intelligence in Education (AIED 2003), 11 pp., Sydney, Australia.

J. Vassileva, R. Deters, J.E. Greer, G.I. McCalla, V. Kumar, and C. Mudgal (1998). A Multi-Agent Architecture for Peer-Help in a University Course. Workshop on Pedagogical Agents, Fourth International Conference on Intelligent Tutoring Systems (ITS'98), San Antonio, Texas, August 1998, pp. 64-68.

Vassileva, J., McCalla, G., and Greer, J. (2003). Multi-Agent Multi-User Modeling in I-Help. J. User Modeling and User-Adapted Interaction 13, 179-210, Kluwer Academic Publishers: Netherlands.

Wiley, D. (2003). Learning Objects: Difficulties and Opportunities. White Paper, The Reusability, Collaboration, and Learning Troupe, Utah State University, Utah. <http://wiley.ed.usu.edu/docs/lo_do.pdf>

Zaizne, O. and Luo, J. (2000). Towards Evaluating Learners' Behavior in a Web-Based Distance Learning Environment. Proceedings of IEEE International Conference on Advanced Learning Technologies (ICALT'2000).