Customizing the Instructional Grid

Beverly Park Woolf, Chris Eliot

Department of Computer Science
University of Massachusetts
140 Governors Drive
Amherst, Massachusetts, U.S.A. 01003

bev@cs.umass.edu; eliot@cs.umass.edu

Abbreviated title: Customizing the Instructional Grid
Customizing the Instructional Grid

Abstract. The Web contains hundreds of thousands of educational resources available anytime and anyplace. However no smart technology is available to help teachers and students locate appropriate resources customized to their needs and social characteristics. When educational resources are indexed, it is often done by demographics, such as student age and grade. This article describes customized Grid Learning Services (GLS) that will personalize instruction based on an individual's presumed knowledge and cognitive and learning needs. The customized GLS will use real-time student modeling, the Semantic Web, intelligent agents and pre-tests of cognitive, affective and social characteristics to personalize selection of educational resources and problems. Components of the customized GLS include an ontology construction agent, goal-based retrieval mechanisms, a lesson planner and student and pedagogical agents.

1 Problems addressed

The World Wide Web is the world’s largest and most flexible repository of education material. The opportunity exists to adapt millions of instructional resources for individual learners, or “mass customization” with respect to education (Davis, 1997). This will help eliminate constraints of space and time in education, and move education from a loosely federated system of state institutions and colleges into a knowledge and learning industry.

However, many instructors and students have difficulty using the Web for instruction. Although schools enjoy greater connectivity and teachers use the Web more (Tech-Ed, 2000), teachers do not utilize the Web for student instruction (Cuban et al, 2001). “Students are frustrated and increasingly dissatisfied by the digital disconnect they are experiencing at school. They cannot conceive of doing schoolwork without Web access and yet they are not being given many opportunities in school to take advantage of the Web.” (Levin & Arafeh, 2002, pg v). Teachers cannot locate appropriate or effective resources on the Web.

Primary concerns about Web instruction are that:
• Students differ in many aspects, e.g. capabilities, cognitive development, goals, requirements and preferences. To fulfill the vision of the Web as the main repository for the way computers and people work together, the individual student has to be taken into account. The current one-size-fits-all-instructional approach to Web resources is outdated.

• No smart technology exists to help teachers locate appropriate resources. Present day indexes, including metadata, identify resources according to student age and school level. No method exists to select resources by student interests or pedagogical characteristics, e.g., an individual’s cognitive level or behavior attributes.

• Educational resources should be available based on pedagogical variables and evaluated based on quality, e.g., through reliable reviews, user comments and ratings.

The quality of Web resources is also of critical concern: enabling full access to poor or non-existing online resources does not provide good education. Currently, no gatekeeper evaluates Web resources as do publishers. No “Amazon.com for education” considers each resource to determine the value of the knowledge source and to communicate with teachers and students about learning goals and appropriate resources.

Currently Web resources are indexed by metadata that focus on the demographics of the resource (e.g., author, student age, grade, media type). Other dimensions important to teachers and learners are not recorded, including social characteristics (gender, main language or ESL learners, ethnicity), cognitive characteristics (cognitive development, spatial ability, math fact retrieval speed), reading level or affective characteristics (self-efficacy, motivation, beliefs-attitudes towards the subject). Nor are features included such as compliance with standards or subject emphasis (e.g., sports or music). Search and browse facilities only solve part of the problem; conventional search engines are designed to generalize and thus to retrieve many matches, see Table 1. They provide one-shot solutions and do not reason about a student’s learning needs, nor evaluate resources based on monitoring student use of resources over time.

The state-of-the art with regard to personalizing content for the Learning Grid is that several systems are being built to customize activities for learners. For example, DIOGENE, an European
Commission project, will support learners within self-adaptive courses and provide personalized self-adaptive courses, using content from registered providers' servers and freeware content drawn directly from the Web, in the dynamic course composition process. This project will also use the Semantic Web to dynamically compose courses. Many researchers have proposed to organize or personalize resources for learners; however, very few infrastructures have actually been built. For example, an assessment framework using agents to diagnose student capabilities has been proposed (Vassileva et al., 1999) as has building agent-based collaborative learning systems. Educational data bases have been built, but these are not Grid systems since they reside on a single machine and are static databases (material is not upgraded, e.g., to remove bugs, unless authors are tracked down and induced to cooperate). They also have limited value as they do not include pedagogical knowledge about their own resources or student learning styles, do not include intelligent searches and are often specialized in terms of media, i.e., not open to the full educational repertoire (e.g., intelligent tutors, animations, multimedia resources, problems, homework assignments or slides). Reusable learning objects have also been proposed as part of the solution, but methods have not been articulated for the automatic evaluation and organization of those objects. Interoperability facilitates resource combination, but does not qualify the resource in the first place.

We propose Customized Grid Learning Services (GLS) to provide learning approaches individualized for each student. This Learning Grid is not just use of the Web and formed from a cluster of static computers, physically contained in the same or fixed locations, although the description below begins that way. Rather, this article describes heterogeneous resources, integrating storage, networking and services. Resources include machines running various operating systems and including the ability to manage the workload. The advantage and challenge of such a Learning Grid is that it is dynamic (resources appear and disappear on the net), inherently distributed, can be located anywhere and offers increased scalability. It is also based on a community of people (from several universities and sites, with different software resources on different web pages). The services are significantly centered on the exchange, the negotiation,

1 http://www.diogene.org/
2 Merlot, National Engineering Education Delivery System (NEEDS) and Math Forum.
3 Collections of reusable resources include EOE (www.eoe.org), NEEDS (www.needs.org) and ESCOT (www.escot.org).
the dialogue within and among virtual learning communities. This Learning Grid will ultimately include a widespread and diverse collections of CPU resources, data resources organized into a virtual file system and learning applications (in this case mathematics problems) organized into standardized, reusable libraries of components. It will support the collaboration of different groups and enable users to collect and organize disparate resources into a more uniform, manageable, visual whole, making the virtual Grid accessible to multiple users simultaneously.

The Customized GLS will maintain a learner model and an electronic vita for each student (respecting privacy requirements) and use metadata and an ontology for knowledge manipulation and intelligent course tailoring. It will use Semantic Web techniques and autonomous agents to develop creditable curriculum for each student and will suggest alternative learning approaches, negotiate customized lessons, construct materials to be presented and will automatically streamline the search, customization and assembly of educational resources. Metadata will be automatically generated from encoded descriptions of instructional resources and the structure and classification of student needs will also be automatically derived.

2 Customization of instructional resources

Research to personalize computer activities has been accomplished within the user-modeling community where user models help tailor a system's behavior to the needs of individuals. However, the Web introduces new challenges brought about by an even larger and more diverse population. Additionally, the Web provides an enormous set of resources as well as new opportunities to learn about the user by tracking the users’ activities to discover patterns of activities. One research issue is to develop and manage an instructional infrastructure that will individualize resources on behalf of all students and be itself modular and extendable.

We propose Customized Grid Learning Services (GLS) to personalize instruction based on an individual's presumed knowledge and cognitive and learning needs. The Customized GLS will adopt well-defined standards and ontologies to provide extensibility, flexibility, interoperability and reusability. This section outlines the goals and objectives of the Customized GLS and the next section describes its components.
The goals of the Customized GLS are directed at the several stakeholders of education and will:

- enable students and teachers to access and effectively utilize resources customized for learning styles;
- encourage designers of instructional resources to provide materials for the Web that teach effectively and contain intellectual merit;
- and provide computer scientists with the mechanisms to develop instructional resources to be used simultaneously by hundreds of thousands of users.

This Customized Grid solution will find the most up-to-date and well debugged material, will handle proprietary and copyrighted material, and will include better materials than can any single static database.

The Customized Grid is not seen as a cure-all for the well-known problems of conventional e-learning. For example, it does not address the isolation experienced by some students studying on the Web, nor the need for live Web mentors. It does provide a way to address the one-size-fits-all production of instructional materials and the need to personalize education for a diverse and heterogeneous population of students.

The objectives are to develop:

- Semantic Web descriptors for resources and student characteristics;
- Intelligent agents that search several digital libraries for appropriate resources;
- Semantic Web descriptors to categorize educational content along dimensions that are important to teachers and learners, e.g., relation to state educational standards, literacy, affective and cognitive characteristics;
- A stable version of the customized environment within several digital libraries; and
- A plan to evaluate the impact of providing customized problems to students and teachers.

We are using extended information about a student's varying patterns of cognitive, motivational and social profiles to customize mathematics problems and will consider the following learner characteristics:

- Cognitive characteristics (cognitive development, spatial ability, math fact retrieval);
- Social characteristics (gender, main language --ESL learners-- , ethnicity);
- Affective characteristics (self-efficacy, motivation, beliefs-attitudes towards the subject).

The objectives are general and the methodology can be applied to any Web collection. The technical plausibility of the proposed scheme needs to be demonstrated before its educational and
economic feasibility are fully understood. Long term goals for the fully Customized GLS include solution of more high level education problems, including: the Classification Problem, in which agents automatically and regularly search the Web to locate, organize and customize educational resources; the Aggregation Problem, in which agents assemble resources to create learning packages, e.g. assemble homework, quizzes and slides of the same topic; and the Evaluation Problem, in which teachers, students and agents assess each resource and verify their usefulness.

We propose that when educational Web resources are customized according to learner characteristics, learning will be vastly improved. To test this hypothesis, we describe an experiment to measure whether customized teaching is more effective than teaching with resources that have been chosen randomly from a databases without consideration of pedagogical dimensions. We propose to use a popular mathematics digital library, Math Forum, supporting one of the largest Web communities (over a million individuals) and most popular instructional digital libraries in existence. The extensive web-site indexes over 1.2 million learning resources,\(^4\) containing hundreds of arithmetic and geometry problems. This library will be integrated with other mathematics tutors on separate servers that also individualized the problems and hints.

The experiment to evaluate whether customized tailored instruction is better than random instruction is admittedly based on an instructionalist teaching approach, in which problems have a singular solution (in this case mathematics and geometry problems). Customizing problems and hints in mathematics, science and engineering academic disciplines is still easier to implement than customizing material within a constructivist, open discovery or situated learning environment (such as in Woolf et al., 2005; White et al, 1999 or Murray et al., 2003). Evaluation of constructivist environments is still extremely difficult (Murray et al., 2005) as students collect data or evidence and construct their own hypotheses and inferences. Thus the tutor is managing disparate and idiosyncratic reasoning on the part of each student. Constructivist tutors will also benefit from customization, which will organize cases for the individual learner and coach each based on evolving cases, student learning needs and social and affective characteristics. However,

due to the youth of the field of Learning Grids, this first experiment will be conducted with a wide community of students and customized mathematics problems residing on remote and heterogeneous computers.

One of the remote resources is Wayang Outpost a geometry tutor that helps students learn to solve geometry problems typical of those on high stakes achievement tests. The student is presented with a multiplicity of activities, including a battery of SAT-Math problems, Figure 1. If the student answers incorrectly, or requests help, multimedia explanations provide step-by-step instruction and guidance in the form of Flash animations with audio. Two forms of help are available: analytic hints that provide an algebraic solution and visual hints that provide animated graphics, such as an angle with a known value rotates and moves over to the corresponding angle with an unknown value. Explanations and hints provided in Wayang Outpost can be tailored to the student’s needs and skills, and therefore resemble what a human teacher might provide, e.g., by drawing, pointing, highlighting critical parts of geometry figures, and talking, rather than by heavy reliance on screen-based text. Individual student skills and learning needs are determined though observation of the student’s behavior while problem solving and through pre-tests, including a standard assessment of mental rotation skill, Figure 2, that measures spatial ability (Casey et al., 1997; Vanderberg et al., 1978). A separate on-line assessment of the student's proficiency with math facts indicates the degree of fluency (accuracy and speed) of arithmetic computation (Royer et al., 1999).

Evaluation of this Geometry Tutor indicated that students who received the visual hints were more engaged by the help than those who received the algebraic hints. Students who received the visual hints asked for more hints and interacted more with the help function than students in the analytic condition (Arroyo et al., 2004; Woolf et al., 2005. Perhaps students were more attracted to the visual help since its strategies are not often presented in the classroom.

The Customized GLS experiment will provide multiple hints and help for each resource, e.g, both visual and computational hints for the Geometry Tutor. The experimental group will use the enhanced and adaptive problems, work on problems every week and be mentored by volunteers.
Each student’s learning style will be monitored and instruction adapted to the individual learner. The control group will work on an equal number of problems, yet hints, and problems will not be adapted to the student’s cognitive or spatial ability.

The Customized GLS will estimate a student’s cognitive, affective and social characteristics, based on pre-tests and real-time students models, e.g., Figure 2. Many intelligent tutors customize content for individual students and modify their behavior based on student models. They are adaptive so that that presentation of each topic is motivated by knowledge of student skills, including adaptive content, and selection of appropriate resources. Content and instructional strategies have been customized to the needs of individual students, varying the pace of instruction and presenting problems in such a way as to challenge students at appropriate levels (Beck and Woolf, 1998; Beck et al., 2000, Eliot & Woolf, 1996; Arroyo et al., 2003; 2004).

3 Components of the Customized Grid Learning System

A detailed solution to the problems inherent in the design and implementation of a Customized GLS is described in this section. We will integrate the power of the Semantic Web defined by W3C, using XML tags to describe digital instructional contents in a machine intelligible fashion. Dozens of evaluated pre-and posts test will be used, see Table 2 that measure student characteristics, such as the role of motivation and cognitive development in learning. Most tests have been validated and used with hundreds of students (Arroyo and Woolf, 2001; Arroyo et al., 2003; 2004). A teacher or learner can select topics from an ontology and the system will personalize (based on learner profiling) a course of material. Student modeling techniques, developed for several mathematics tutors and tested with nearly 1,000 students, can assess students with varying levels of cognitive and motivational processes. An autonomous model of the student will communicate with an autonomous model of instructional resource, Figure 3. A suite of tools, Table 3, will automatically streamline the search and customization of resources for a restricted set of library material.
We have designed a multi-agent system (MAS) to perform information gathering and sophisticated reasoning in support of context sensitive planning to sequence educational resources, based on specific student needs and feedback and have developed autonomous functionally specific agents, each specialized at solving a particular problem aspect. In the Web, information sources, communication links and agents can appear and disappear unexpectedly. Thus, the educational agents need a sophisticated reasoning architecture to operate asynchronously to address many difficult challenges and achieve the goals stated above.

Currently, we are working on four components of this solution:

3.1 The Learning Semantic Web
3.2 Automatic ontology construction agent.
3.3 Goal-based retrieval mechanisms.
3.4 Lesson planner.
3.5 Student and pedagogical agents.

3.1 The Learning Semantic Web

The Semantic Web (SW) is still very much at a grassroots level: People are starting to publish information using the Resource Description Framework (RDF) developed by Tim Berners-Lee and thereby making their databases and applications fit for the Semantic Web. OWL might be used instead of RDF as it outperforms the latter. Any one can create a Semantic Web application, simply by publishing an RDF that describes a set of URLs, what they do and how they should be used.

We define the Learning SW as a mesh of instructional resources linked in such a way as to be easily processable by machines, on a global scale. This is an efficient way of representing learning resources on the Web, or as a globally linked database. Although the basic parts of the Learning SW, RDF and the concepts behind it are very minimally constraining, applications built on top of the Learning SW will be designed to perform specific pedagogical tasks, and as such will be very well defined. People need to publish their own Learning RDFs and then other people’s Learning RDFs can link to them. In addition, there is no point in reinventing the wheel; viz., if someone has already invented a learning schema that contains comprehensive and well-understood entities, then others should adopt it. We need a grassroots effort to link instructional resources together and encourage people to use other people's terms effectively; terms such as...
"dc:title" mean that a Dublin Core application could "understand" a person’s code if information from one application is to be repurposed to another resource the near future.

Three sizes of Learning SW are emerging. A small scale Learning SW is a set of explicitly written learning resources, e.g., collections of slides, examples, or problems on a given topic located on a few computers. This may include a dozen courses on a specific topic, e.g., learning information technology, that will be used primarily by a known group of users, or applications that will be transferred, perhaps between learners within several classes or organizations. A medium scale Learning SW attempts to make sense out of several small-scale Learning SWs joined together. An example is the integration of two previously established learning systems, e.g., several databases of chemistry or physics problems. This requires writing a Learning RDF that integrates the features of both learning systems. Large scale Learning SW involves several large databases of learning materials and heavy duty inference rules and processors to handle the learning material. An example of a large scale Learning System is several organizations that merge all their resources, say hundreds of learning activities for teaching mathematics.

We propose a Resource Description Framework (RDF) for a medium Learning SW that integrates several databases of tutors.

**RDF Schema for Teaching Resources**

:Teachingtutor rdf:type rdfs:class
:TeachingtutorTitle rdf:type rdf:property
:TeachingtutorTitle rdfs:domain :Teachingtutor
:GeometryTutor rdf:type :Teachingtutor
:GeometryTutor rdf:tutorTitle “GeometryTutor”

**RDF Schema for Student Characteristics**

:CogDevelopment rdf:type rdfs:class
:CogDevelopmentHigh rdf:type rdf:property
:CogDevelopmentHigh rdfs:domain :CogDevelopment
:CogDevelopmentLow rdf:type rdf:property
The RDF schema above simply says that Teaching Tutors are a class, with Teaching Tutor Titles, a property of that class, restricted to the domain of Teaching Tutors. The Geometry Tutor is a type of Teaching Tutor. The RDF schema for student characteristics says that cognitive development and spatial ability are two classes of data (for students) and that they can range from high to low for an individual. The next step is to make inferences about the stored learning resources, or to derive new data from resource and student data already known. In the example above, the GeometryTutor in the resource database, can tutor by focusing on its spatial or computational hints. The spatial version is appropriate for a student with SpatialAbilityHigh, whereas the computational version is more appropriate for a student with SpatialAbilityLow. Unfortunately, great levels of inference can only be provided using "First Order Predicate Logic" languages, and, current Semantic Web inference languages, such as DAML, are not a FOPL language entirely. However, agent technology is well suited to make inferences about the student, as suggested here and described in the next section.

3.2 Automatic ontology construction agent

The Learning Semantic Web technology will be integrated with agent technology that will handle inferences about the student’s learning variables and rich instructional resources. Agents perform information gathering and sophisticated reasoning in support of planning to sequence educational resources, based on student need and feedback (Woollf, 2000; Cassin et al., 2003; 2004). Each agent specializes in solving a particular problem, e.g., student support or indexed resources. Agents improve their problem solving ability through communication and coordination. By planning the content, amount, type, and timing of their communication (Sycara, 1998) or by using
abstraction and meta-level information (e.g., organizational knowledge), they decrease communication overhead (Durfee, et al, 1987).

We developed an *Ontology Agent* to automatically construct an ontology of online course descriptions (Cassin et al., 2003; 2004). The goal of this agent was to identify all instructional resources for a specific topic, say introductory programming. This agent understood the content and approach of a teaching resource. We used information retrieval methods to summarize word usage as high dimensional vectors to extract online catalogs for 100 courses on the topic of computer science from 40 universities. Using only descriptions and syllabi, we applied word frequency vector cluster algorithms to determine which courses were similar, by comparing their word vectors. The results were extremely accurate; each cluster contained a pure topic, e.g., introductory programming, artificial intelligence, etc. Subtopics of these structured documents and course curricula were compared to determine the common kernel of subtopics and the documents merged by averaging the word frequency vectors. A *pedagogical topic* (or simply topic) is the basic unit of the ontology. The granularity or level of abstractness can be quite varied and range from the topic of “computer science,” down to the granularity of “linked list” or “pointer.”

This Ontology Agent used information retrieval methods on Web course descriptions to summarize word usage as a vector in a high dimensional space. We assume there is a regularity to the order in which subjects are taught (Cassin et al., 2004). In most cases, this regularity represents knowledge gained from experience teaching the subject, generally thousands of times and possibly to millions of students. There may be alternatives that work, but there are certainly learning sequences that do not work. Evidence for these orderings is easily found, since most course materials have some linear organization. The reasons for choosing one learning sequence over another may be explained by psychology, or in domain specific terms, or because of institutional or economic considerations, but we are not concerned with justification. We assume that existing educational material is (reasonably) well-designed and our goal is to determine the structure represented by existing best practices. We determine the pair wise frequency with which
one topic appears before or after another and assume that a prerequisite relationship exists, when one ordering is more prevalent than others.\(^5\)

New instructional resources appear regularly on the Web. When one is discovered, the challenge is to recognize what material it covers and whether it is strongly related to a previously defined topic, but still distinct from it. This is similar in spirit to the problem of “new event detection” in news analysis, where a system recognizes that a story talks about something that had not been discussed in the news before (Allan et al., 1998; Yang et al., 1998). We believe that the training and threshold approaches used in that task can be extended to work in this domain also. The Ontology Agent might locate the pre-conditions and post-conditions of the resource assuming this is available in the metadata. The long-range goal is to develop methods that can maintain the ontology knowledge autonomously when needed or with human supervision when desired. Automatic ontology maintenance involves matching subtopics and determining their relationships. Eventually, complete libraries will be searched and analyzed to obtain ontology and planning knowledge.

### 3.3 Goal-based retrieval mechanisms

Once automatic ontologies of Web resources are available for each academic domain, the next step is to dynamically generate several instructional resources for an individual student. This step is divided into a planning and scheduling phase. In the planning phase, a plan network will be generated that encodes a large number of possible solutions to the educational request. In the scheduling phase, a specific schedule will be extracted and optimized. An existing scheduler, Design-to-Criteria (DTC) (Wagner & Lesser, 1999) produces a current “best” solution and then enables a student or teacher to alter the parameters to retrieve a second solution. Given the constraints supplied by the student, DTC offers a variety of “best” solutions. Prior work in building student models and developing algorithms enables us to integrate information about each student’s current and past performance and to predict performance of resources.

\(^5\) A more sophisticated approach is to assume that the observed educational sequences are products of some general model (such as a Markov model) and use machine learning methods to parameterize the model to match the observed educational sequences.
The plan generated in the *planning phase* is input to the scheduler, which produces a comprehensive linear instantiation of one possible solution to the problem, based on constraints, such as preferred time, quality or cost of the teaching materials. Complex goals will be decomposed into simpler sub-goals that may be independently satisfied by various resources. The resources that most helped prior students will be given a high priority. If the student needs to learn something before an exercise can be attempted, then additional learning goals will be generated automatically. The planner can create educational plans optimized for factors such as cost, scheduling in time and space, cognitive style and personal preferences in addition to prior knowledge and desired knowledge.

### 3.4 Lesson Planner

Once instructional resources are accumulated for a particular student, the next step is to organize these resources for that individual’s particular learning needs. Specifically, the *Lesson Planner* customizes the available educational resources. First it attempts to satisfy open sub-goals, such as the trivial sub-goal illustrated in Figure 4: “Does the student know integration by parts?” The *Lesson Planner* will retrieve resources while applying constraints and quality evaluations. When a consistent and complete plan is found, it is presented to the student or teacher for approval. Further searches may yield a more optimal plan, by exploring more options or as a result of unexpected changes in the availability or quality of resources. The *Lesson Planner* engages in a heuristic search for a consistent plan, e.g., resources relevant to the topic and hints and help appropriate for the student.

A student must satisfy all prerequisites before starting to use a learning resource. Student models derived by the *Student Agent* (see next section) will accommodate students with a wide range of skills; students who are strong in one skill can move ahead quickly and others obtain extra assistance in areas where needed. For example a student may have mastered biology, have a very good understanding of the English language and a poor grasp of engineering. The *Lesson Planner* will know enough about the student to gather resources together into a coherent whole and uses well known planning techniques, e.g., (Eliot & Woolf, 1996), possibly in the form of IF-THEN rules, to construct a chain of operators that transform the world from its current state into a
desired goal state. Automatic planners have long been used for applications such as business management systems and robotic problem solvers.

The curriculum-planning problem is much simpler than the general planning problem addressed by sophisticated planners. General planning systems require sophisticated mechanisms to deal with negative post-conditions, for example, correcting a misunderstanding may be part of the plan. As a result, the curriculum-planning problem is comparatively simple, from a computational perspective. Each learning component will be described using an IF-THEN rule, see Figure 5. We will determine which student records will be used in course assembly and how can plans be most clearly presented to students and teachers. Planning will be automatically extracted from an existing site, with some planning decisions automated and some supervised by humans. The planner must formalize some subtle learning goals, while individual learning styles and disabilities will constrain planning decisions.

3.5 Student and Pedagogical Agents

Once the Customized GLS has selected appropriate resources for an individual student and a plan is provided, the next step is to ensure that each student receives and is engaged in the appropriate resource and that each resource is fine-tuned for that student’s learning needs.

Student Agents (SA) are responsible for making the Customized GLS interaction smooth and effective (Figure 6, left). They buffer malicious or sub-par performance by resources (e.g, a buggy application) and assemble various educational resources together into as coherent a whole as possible. A SA contains information about its student's educational background and preferences, including a sparse version of the ontology, with nodes present for those topics about which the SA has a record and solicits student input to ensure that the right choices are being made. A common thread running through the operation of the SA is that of collaboration with the student. A student profile captures student characteristics after only a few student interactions, using machine learning and a record of hundreds of prior students using the resource (Beck et al., 2000). We have develop reinforcement learning, Bayesian network and decision tree models of students to characterize learner features such as known and unknown skills, engagement,
motivation and the likelihood that students will answer the next question correctly and how long to do so (Beck & Woolf; 1998; Arroyo and Woolf, 2005; Jonsson et al., 2005).

**Pedagogical Agents (PAs)** represent instructional components, with a single agent representing a single topic (Figure 6, right). Thus the PA for “algebra” will negotiate on behalf of all introductory or advanced algebra materials. It will know about all the subtopics of algebra, including which pre-requisites are required for each new topic. In its purest form, the PA is to its educational resources as the SA is to its student: negotiating in its best interests and on the lookout for misbehavior, e.g., a crashing resource or a student unable to do the task. Each PA maintains a model of all of its resources, including Learning SW language, meta-data, and both human and automatic evaluation of the resource. Each resource will be put through a certification process and evaluated by independent experts in the topic area. If education stakeholders provide a stamp of approval, the resource is inserted with a high rating. If the resource does not live up to the rating, it will quickly reach its appropriate level. Thus, the GLS will support designers of instructional resources to develop effective and engaging materials, as they will be selected over and over again and the designer will gain revenue.

PAs might potentially represent planning services instead of educational resources. PAs and SAs will negotiate over customization of a resource as to duration, start/end time, specific topics to be covered, additional support to be included, such as an online human tutor, teaching methods to be used, and so on. Not least, of course, will be the cost to the student of using the educational resource. Depending on the student's preferences, known to the SA and stored in a student model, different customizations will be weighted differently in the final course assembly. PAs will ultimately have learning mechanisms that automatically estimate the optimal sequence and rate of topics for each student (Arroyo et al., 2003). They will estimate and organize the rate and presentation of resources based on the student model, predict the time for a student to use a given resource and then adjust the resource to provide a more optimal package for each individual student.

---

6 Instead of representing a single topic breakdown, the agent might actively parse and search a resource to build up a composite model of the topic. This will be useful in the case of difficult queries for which there are no direct tutors or breakdowns.
4 Conclusions

This research expanded the use of Grid educational resources to customize instructional resources for individual students and thus improve teaching as a function of instructional history. This article described Customized Grid Learning Services (GLS) that will individualize instruction based on reasoning about a student's presumed knowledge and learning needs and will customize resources, problems and hints. The Customized GLS will use real-time student modeling, the Semantic Web, intelligent agents and pre-tests of cognitive, affective and social characteristics to personalize educational resources and problems. We proposed an experiment to test the hypothesis that customized learning is more effective than choosing resources from a fixed database without consideration of pedagogical dimensions. The experiment will determine: 1) Whether automatic customization extracts a reasonable set of arithmetic problems that are organized logically; 2) The amount of modeling required to develop a reasonable student model; and 3) Whether students will work longer at customized digital library sites. Evaluation will assess the quality and usability of retrieved components, checking their interactivity, coverage, authority etc., and rate the student's learning. The components of the customized GLS were described, including an ontology construction agent, goal-based retrieval mechanism, lesson planner and student and pedagogical agents.

5 Acknowledgments

Research on Wayang Outpost, the Geometry Tutor was funded in part by the National Science Foundation, NSF-HRD-0120809, “AnimalWorld: Enhancing High School Women’s Mathematical Competence,” C. Beal, PI, B. Woolf, co-PI. Research on Student and Pedagogical Agents was funded by NSF IIS-9977960, “Multi-Agent Instructional Communities: A Computational and Experimental Approach,” B. Woolf, PI.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the funding agencies.
6 References


Computer Science, Lecture Notes in Artificial Intelligence, van Elst, L., Dignum, V. & Abecker, A. (Eds.).


Technology in Education, Market Data Retrieval, Shelton, Conn.,


Customizing the Instructional Grid

<table>
<thead>
<tr>
<th>National Science Digital Library (NSF)</th>
<th>Math Forum Digital Library</th>
<th>Math Wizz</th>
<th>20,000 Problems under the sea</th>
<th>Eduseek</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over 33,208 links to mathematics materials, for researchers, students and teachers, sorted by source of repository</td>
<td>Over 925,035 indexed learning mathematics resources, problems including tools and descriptions of solutions for teachers.</td>
<td>Thousands of problems and help for K-12.</td>
<td>Searchable database referencing over 20,000 math problems</td>
<td>Thousands of mathematics resources, Sorted by topic and student age.</td>
</tr>
</tbody>
</table>

Table 1. Search engine results for “mathematics problems” as of July 2004.

![Figure 1. Two forms of help in the Geometry Tutor. 1) Visual hints (animated lines on the figure) propose that the student mentally translate angles to determine the missing value; and 2) A fairly traditional computational approach, such equations (right of the figure).](image)

Figure 1. Two forms of help in the Geometry Tutor. 1) Visual hints (animated lines on the figure) propose that the student mentally translate angles to determine the missing value; and 2) A fairly traditional computational approach, such equations (right of the figure).

![Figure 2. Assessing a student’s spatial abilities through an on-line mental rotation test.](image)

Figure 2. Assessing a student’s spatial abilities through an on-line mental rotation test.
Cognitive and Spatial Abilities
- Piagetian Cognitive Development
- Spatial skills pre-test
- Math-fact retrieval pre-test

Affective Characteristics
- Mathematics Attitudes Scales
- Self-confidence/math importance
- Eccles Attitude Test

Domain Skills
- Arithmetic achievement test
- Geometry pre-test
- Mock SAT post-test

Table 2. Internet-based tests to analyze student characteristics.

Table 3. Modules in the Customized Grid

<table>
<thead>
<tr>
<th>Module</th>
<th>Target Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td>* Student Agent</td>
<td>Represent and promote the student; assists in the creation of courses and monitor student performance.</td>
</tr>
<tr>
<td>* Pedagogical Agent</td>
<td>Represent instructional resources and sells services to the Student Agent</td>
</tr>
<tr>
<td>* Ontology Agent</td>
<td>Represent available educational resources by extracting terms, precondition and post-condition information from meta-data.</td>
</tr>
<tr>
<td>* Lesson Planner</td>
<td>Assemble resources, maintain student models and create lesson plans.</td>
</tr>
<tr>
<td>Evaluation Agent</td>
<td>Monitor student learning and integrate human supplied resource evaluations with statistical data.</td>
</tr>
<tr>
<td>Student/Teacher Interface</td>
<td>Accept student learning goals and work with teacher to assemble courses automatically.</td>
</tr>
</tbody>
</table>
Customizing the Instructional Grid

131--Calculus I (R2) 4 credits
Elementary techniques of differentiation and integration of algebraic and trigonometric functions, elementary differential equations. Applications physics, chemistry, and engineering. Students will use calculators on homework and exams. Prerequisites: high school algebra, plane geometry, trigonometry, and analytic geometry.

IF
Knows (High-school-algebra) AND
Knows (Plane-geometry) AND
Knows (Trigonometry) AND
Knows (Analytic-geometry) AND
Passes (MATH-131)
THEN
Knows (Differentiation) AND
Knows (Integration of functions) AND
Knows (Elementary-differential-equations)

Solution: Passes (Pre-Calculus)
Passes (Calculus)

Figure 5. Formalized Course Entry

Goal: Knows ? (Integration by Parts)

RULE-1:
IF
Knows (Algebra) AND
Knows (Geometry) AND
Passes (Pre-Calculus)
THEN
Knows (Functions)

RULE-2
IF
Knows (Functions /Continuity)
Knows (Integration)
Passes (Calculus I)
THEN
Knows (Integration by parts)

Figure 4. A Trivial Lesson Planner Problem
Customizing the Instructional Grid

Figure 6. Functionality of Student Agent (SA) (left) and Pedagogical Agent (PA) (right).