Intelligent Virtual Environments

Intelligent Tutoring Systems for problem solving emerged as a leading success story among first-generation intelligent learning environments. These ITSs effectively support learning-by-doing, across a variety of domains and age groups. This white paper builds on the success of ITSs:

- ITSs for mathematics, science and programming have been shown to yield learning gains that exceed those of the typical human tutor and approach the success of the best human tutors.
- ITSs have been piloted successfully at the college levels and more importantly, ITSs have begun to penetrate real-world classrooms. Approximately 500,000 middle school and high school students are currently using intelligent mathtutors annually.

But this white paper recognizes that the scope of this success is limited in multiple ways:

- ITSs are largely limited to formal content domains (math, science, programming) and military tasks (e.g., equipment operation and troubleshooting; tactical decision making).
- Within these domains ITS technology is quite successful at analyzing and responding to the accuracy of student actions, but is in the early stages of recognizing and responding productively to student intentions and to student affect.
- The technology has had minimal impact on formal education outside of middle school and high school math classes, and little impact on informal educational contexts.
- ITS success is limited to situations in which the software fully defines the learning goals and activities. When problem-solving activities are defined exclusively by the technology, there is an increased risk that students focus on simply completing the problems, rather than treating them as learning opportunities.
- The technology generally relies on conventional instructional activities to prepare students for learning-by-doing. This increases the risk that students engage in shallow reasoning during problem solving, and limits opportunities for remediating shallow knowledge.

In this white paper, we (1) propose key goals for the next generation of intelligent learning environments; (2) outline a learning environment framework to support these goals; and (3) propose grand challenges which emerge from this framework.

<u>Scaling the Technology.</u> Each of the following top-level goals is a key component in scaling the technology, to magnify the contribution of second-generation intelligent learning environments to the learning.

- Expanding the coverage of content domains and expanding the range of interactive learning activities supported by intelligent learning environments (including interactive text and graphics comprehension)
- Accommodating learning goals of multiple stakeholders. First among equals are the student’s learning goals - enabling students to participate more actively in defining and realizing learning goals – but including teachers, parents and educators.
- Scaffolding student goal-setting metacognition. Research on, and support for metacognition, largely focuses on learning how to *achieve* learning goals. Students

increased participation in setting learning goals will benefit from metacognitive knowledge on how to *set* goals successfully.

The ALOE Environment. We envision an Active Learning Object Environment (ALOE) to achieve these goals. ALOE is related to the learning object economy proposed over a decade ago, but with two essential advances:

- *Active* Learning Objects. While the earlier learning object economy, which consisted largely of passive objects, e.g., videos, ALOE will consist of *active* learning objects, which elicit student actions and provide feedback.
- Multiple Stakeholder Learning Management System. The active learning objects will be organized to facilitate student-initiated learning goals, as well as learning goals initiated by other stakeholders.

Active Learning Objects. Active Learning Objects (ALOs) are essentially unit tasks; they are atomic elements which define the content of a curriculum, although completion of an ALO task is further analyzed into steps. An individual problem in an existing intelligent tutoring system is an ALO. As in an existing ITS, an ALO will be capable of providing step-by-step accuracy feedback. Other ALOs will focus on encoding, rather than applying knowledge, but will engage students in active processing of the newly encoded knowledge.

Like many learning environments, ALOE can be characterized by three embedded activity loops:

- The Step Loop. Activities within an ALO occur in a "step loop." An active learning object poses an atomic interactive task in the curriculum, and monitors the student's step-by-step actions, providing advice as needed for students to complete the task.

- The Multi-stakeholder Learning Management System: The Task Loop. Activities of the LMS in designing a curriculum (a sequence of ALOs) occur in a "task loop."

- The Lifecycle Loop. ALO development activities occur within a "lifecycle" loop. These include design, implementation, testing, and dissemination.

Design goals and challenges for the ALOE environment at each of these levels are described in the following sections.

Grand Challenges. Several grand challenges which emerge from this analysis:

- Developing Active Learning Objects capable of participating in learning dialogues with (groups of) students
- Developing an active multi-stakeholder learning management system that helps stakeholders select appropriate active learning objects for students.
- Developing improved lifecycle monitoring and management capabilities.

These grand challenges are expanded in the following sections.

Step loop challenges

A task (supervised by an ALO) is made up of a number of steps and, in general, takes about 20 min to 2 hours in one session. Each step can take a few seconds, and there can be up to or more than hundreds of steps per task/activity. Steps may be laid out in space (e.g., as parts of an essay or a schematic drawing) rather than time. An active learning object (ALO)

- monitors the student's activities step-by-step,
- understands what opportunities exist in that activity for improvement (relative to stakeholder goals), and
- plans/executes ways to get the learner to take advantage of those learning opportunities.

Because these capabilities each have a number of challenges, they are described in 3 subsections that follow.

Monitoring activity and thinking

Monitoring is the capability to take sensor data and interpret (e.g., categorize) it without judging its "correctness" but merely understanding it in terms of progress toward end states. Challenges exist in monitoring progress of activity process, including such things as progress, completion, and learner thinking. (before data mining?)

This component of active learning objects will be discussed in terms of the various sensor modalities, e.g., speech vs. text. The current state of the art is briefly presented along with challenges.

Input Modalities: Monitoring Student Behavior.

Structured form-based input is the most highly developed input modality (e.g., filling in an html form), which is reliable and accurate by design. There are good tools for creating forms and their interpreters. Examples of active learning objects that use form-based input include the cognitive tutors, CTAT tutors, the SQL tutor, and many others. Challenges:

- (See cross-cutting challenges below)

Canvas-based input is almost as highly developed. It differs from form-based input in that the location of the user's input is decoupled from its expected semantics. Students can enter text, graphical gestures, etc. anywhere on the canvas. Because the interpreter does not know what "should" be entered at that location or at that time, it must base its interpretation solely on what the student has entered. This can be quite difficult if the student has entered something unexpected, such as an incorrect step. In this case, the interpreter may have a difficulty figuring out what the user "should" have entered.

- (See cross-cutting challenges below)

Free-text understanding is coming along but needs to become more reliable and accurate. This includes understanding the students' typed-input turns during dialogue, either amongst themselves or with a computer. It also includes understanding essays, reports, long explanations, and other monologues that have long, substantive content. Both dialogue and monologue understanding have benefited from advances in statistical language processing. The challenges include cross-cutting challenges listed below plus:

- Taking advantage of existing linguistic resources e.g., WordNet, Comlex, etc.
- Reducing the amount of data needed for calibration of statistical recognition/interpretation

Free-speech understanding is proceeding more slowly. It is reliable now only in dialogues where the computer asks short-answer questions or other questions where only a few types of easily detected responses are expected. Improving continuous speech recognition is important for many applications, but the educational application has certain advantages that are currently not being seized. For instance, in some instructional applications, the computer doesn't need to recognize everything being said, but only the things that it can interpret as learning opportunities. Or, it may only need to detect who is talking, their affect (via prosody) and length and affect of the responses. Challenges (see also cross-cutting challenges):

- Near term: Develop work-arounds that let the system interpret free speech without completely understanding it.
- Longer term: Just better speech understanding.

Sensors embedded in actual equipment (as with the AEGIS embedded trainers). Support of learning-by-doing, requires monitoring tool use, both in the cyber space, and for performing physical tasks (e.g., gauges; drills; etc.).

Physiological sensors, movement monitoring, video camera feeds of student and/or their environment (see Interface group).

Ultimately, the goal is to be able to make sense of the student as a human tutor or peer would, by seeing what the student is doing and saying, and evaluating it in terms of finding learning opportunities, collaborating, etc.

Cross Cutting Challenges: Interpreting Student Behavior.

The challenges in identifying students' behaviors described above are just the first step toward recognizing and acting upon learning opportunities. The following cross-cutting challenges emerge in interpreting student behavior:

Interpreting Student Responses. Even when a student action has been successfully identified, there can be ambiguity about how it relates to the overall solution structure, and whether the action is correct. This in turn interferes with providing suitable feedback and recognizing learning opportunities. Sources of ambiguity about an apparently correct action include:

- a student may respond correctly based on deep domain understanding or based on superficial reasoning strategies;
- a student may respond correctly by guessing;
- a given action may be correct for more than one task subgoal, and it can be unclear which subgoal the student is working on;
- a given action may map correctly to one task subgoal, but it may be incorrect for the task subgoal the student is currently working on.

Challenges, mostly achievable in the short term, include (most of these have been realized in pilot research, but all need to be further developed to achieve greater accuracy and to be extended across domains):

- Using activity-state context to disambiguate
- Preserving ambiguous interpretations for as long as possible
- Using population priors to guess interpretations
- Using data mining to determine new approaches, paths, and behaviors in problem solving (maybe more appropriate for open-ended and discovery systems, rather than tutors per se).
- Successful "guess" detectors have been developed in math problem solving tutors, but need to be extended to other domains.
- Deep and shallow reasoning detectors.
- Detecting the nature of activity pauses (on-task thinking, on-task help-seeking, or off-task behaviors)
- Interpreting the intent of student turns in dialogue and group collaborations.

Detecting Student Affect.

Challenges, mostly achievable in the short term include detecting the following, both from parameters of students' problem solving actions and from monitors of the students themselves:
- Interest (intrinsic motivation)
- Boredom
- Frustration

Finding learning opportunities

Suppose all the data are being collected and interpreted up to a certain level; the ALO knows what the student is doing, so now it's time for judgment: when and where is there room for improvement? What kinds of learning opportunities can the ALO find? This is not just about correcting errors or guiding students to follow efficient paths to get the correct answers (perhaps "long tail" approaches to problem solving), but also about noting opportunities to improve affect (motivation, interest, emotions, self-efficacy), "21st century skills" (teamwork, leadership, critical thinking, communication skills, etc.), learning styles (one kid likes to explore, another likes to follow directions; one likes video and visuals, another likes text; curiosity, focus). Although these analyses feed directly into assessment (of the learner) and data-mining (for evaluation of the ALO or treatment), we focus here on learning opportunities that trigger changes in the course of the ALO's interaction with this learner during this (half-hour) period of activity. Learning opportunities also include excellence that needs to be reinforced, ingredients that will be used later in learning events ("just remember this, because we're going to discuss it later") and things that should be interesting to the student.

Challenge	**2 yr**	**20 yr**
How to determine what is "correct" in view of the stakeholders (e.g.,		x

instructors), especially in open-ended environments that provide many “correct” paths, answers, etc.		
Comparing “correct” to students’ path	x	x
Classifying differences between student and “correct” paths in terms of predicted eventual outcomes (e.g., serious misconceptions vs. minor mistakes vs. missed opportunity for interesting sidetrack….)	x	x
Detecting gaming the system; bad gaming vs. good gaming	x	x
Finding opportunities even when system doesn’t understand the user’s ultimate intentions during the activity, and there is no “correct” path.		x
Measuring affect. How to identify eureka moments, or moments of intrinsic interest or frustration.	x	x
Measuring 21st century skills. E.g., what is good teamwork, how can you tell? Communication skills. Good peer dialogues.	x	x
Opportunities to practice a meta-skill, e.g., checking one’s work before submitting; self-explaining a bottom-out hint; etc.	x	

Taking advantage of learning opportunities

Once the ALO has located some learning opportunities, it must decide what to do about them. That is, what should it do, if anything, that will get the student to seize the learning opportunity and perhaps even learn from it? This includes determining what kinds of feedback and adaptivity are appropriate, what kinds of direct instruction if any, and what learning objectives to send off to other ALOs that will be run later by the student.

Challenges	**2 yr**	**20 yr**
Matching the sensory data with the student profiles to determine a best approach.	X	x
When: before, during or after the activity or at all	X	x
What types of recommendations to make at the end of an activity (e.g,, what to do next: If a student shows an interest in X, and many people who like X also find Y interesting, then suggest Y. Or in response to displayed weakness or challenge.)	X	x
Match feedback to motivational style of the student. (e.g., stars, competition,	X	x

acknowledged by peers, get attention, perform, personal improvement).		
Manipulating or adapting the game/scenario/ALO/ environment.	x	
Awareness of pre-requisite and follow-on activities. This is the map of the content area (just the content area?).	x	x
Becoming a coach, doing such things as facilitating communities of learning, including peer feedback (what do you think, John?). (research question: does the coach have to understand the content area, or just social conventions and metacognitive scaffolds?)	x	x
What: didactic vs. discovering corrective info by oneself	x	
Insuring that feedback is constructive (gets students to learn) rather than discouraging (convinces student that student is “not good” at the activity)	x	x
Insuring stakeholder (especially the student) control. E.g., “I’m really into this right now; hold your advice for later.” Or “Ouch. I resent the way you gave me advice.”		
Self-improvement; ALO’s policies on when/how to give advice change as it sees whether students learn from the advice.		
Customizing feedback and coaching techniques to student traits such as personality, learning style, motivation, and culture, and to student states such as affect, level of engagement, level of frustration, etc. A number of studies in the field of cognitive psychology have documented the benefits of different types of feedback techniques (e.g. immediate feedback vs. delayed feedback). Recently, data mining techniques are being used to study the effectiveness of different types of feedback and hinting techniques to different student characteristics. It is expected that there are several low-hanging fruits that can be exploited in the immediate term. However, there is much potential for long-term exploration.	X	X
Matching a learner with other learners and/or mentors taking into account learner models and interests. This is in support of feedback and coaching provided by peers and mentors.	X	
Seeking and presenting just-in-time references that take into account the learner’s current task context, prior knowledge and mastery and preferences.	x	

The task loop: Selecting an active learning object

Embodied in ALOE is a vision for an active multi-stakeholder learning management system (AMS-LMS). "Multi-stakeholder" refers to the idea that all the relevant stakeholders (e.g., students, mentors, teachers, parents) express their goals and some kind of negotiation occurs for the sake of the individual learner. "Active" refers to the idea that the system is part of this negotiation, offering advice, theory and data. The AMS-LMS also orchestrates this negotiation and implements the resulting policy. It stores learner profiles, stakeholder concerns and many other data required for achieving its function.

Whether the AMS-LMS actually stores the ALOs, as do current LMSs, is an open question. Our preference is that content be scattered about the web, but tagged so that ALOE can find it and consider selecting it. To be considered for selection, the ALO must of course have metadata describing its content and its prerequisites. They should also describe their pedagogy. For instance, there might be different ALOs that address the same concept but use different pedagogical approaches such as "teaching by demonstration", "teaching through inquiry", "teaching with Socratic dialogs", or "collaborative learning." A description of the pedagogy is included in the ALO's meta-data. ALOs may also embed human-readable instructions to the student, group or instructor on how they are to be used. Similarly, an object's metadata include descriptions of modalities, cultural biases, and any other information that may be useful for selecting the ALO.

Although the stakeholders ultimately set policy and make decisions about which ALO to select, they often want to know how those selections will affect the student's future. Thus, the AMS-LMS should be able to make predictions about: (a) areas in which individuals have the proclivity to master, perhaps including career potential in relation to learner interests; (b) future performance in skill areas; (c) time to master a skill area or become an expert (d) topics that the student should find interesting, (e) given a particular topic/ALO, students who would work well together in a group. These predictions are driven by data-mining, recommender system logic, metric models of learning, etc.

Challenges in developing the AMS-LMS

Grounding the meta-data: A major, familiar challenge is insuring that metadata descriptions mean the same thing when used with different ALOs. One approach is to define central ontologies of learning objectives that are used to organize and index the ALOs. An alternative is the folksonomy approach where structure emerges from decentralized tagging. Both these approaches are in use and their relative merits are already being evaluated. Bringing the results of these analyses into the ALO research and recommending an ALO architecture is an important challenge (time frame: current)

Embedded assessments: In order for the AMS-LMS to maintain a model of the student, every (or at least, many) of the ALOs must conduct some kind of embedded assessment relative to the descriptive terms of the metadata. For instance, if the student has just conducted a flawless negotiation, which descriptors are upgraded? Also, if an ALO notices a learning opportunity that it cannot handle (e.g., the student is weak on a certain pre-requisite), then it needs to report that out to the AMS-LMS so that it can suggest making that weakness a priority. (time frame: 20 years)

Orchestrating negotiation: It is all well and good to say that stakeholders will negotiate about selection of an ALO or sequence of ALOs, but how will the system help this occur? How much training and help will the stakeholders need? What kinds of user interfaces? How can they access the predictions? How can they interpret the current model/profile of the student, which exposes not only competencies but many other data as well? (time frame: 20 years)

Informing the negotiation: Although there are certain well-accepted policies for selecting ALOs (e.g., keep students in their ZPD; model-scaffold-fade), how can evidence from the learning sciences be brought to bear on the decision making? Would it only affect the predictions, and only via them would theory influence decision making? If so, how could the stakeholders find out the evidence or warrants behind a prediction? (time frame: 20 years)

Calibrating the students' self-monitoring: Given that the AMS-LMS can make predictions about the students' interest, competence, etc, the system can support growth in the student's self-monitoring and self-efficacy judgments. For instance, it can ask learners to make predictions about their own performance, and then giving feedback and recommendations based on actual performance. (Time frame: 20 years)

Discovering implicit goals: It is possible that stakeholders, especially students, have goals that they are not aware of or have misdescribed. In this case, their choices of ALOs may exhibit a pattern that can be recognized and brought ot the student's attention. (time frame: long term)

Lifecycle loop:

The active learning objects (ALOs) that constitute the core components of the envisioned Intelligent Virtual Environment are not static, but rather are dynamic involving a lifecycle consisting of design, creation, testing, acceptance, modification, and eventually perhaps destruction and recycling. ALOs exist in the context of communities that would shape the requirements and place constraints on the lifecycle in order to make the ALOs fit the communities' needs. The ALOs will need to co-evolve with needs, understandings, and the cultural context of the community.

For ALOs to be adopted, the community must trust them and this has several components:

1. The community trusts that the ALOs fill their needs.
2. The community trusts that the ALOs fit within their culture.
3. The community trusts they have the means to evaluate the pedigree, intent, and authenticity of ALOs.
4. The community has a good understanding of the ALOs, what they do, and how they do it.
5. The community trusts that the ALOs do what they are designed to do.

6. The community needs to be confident that the ALOs have good privacy and security policies in place.
7. The community has a stake in the educational process and can exercise control over the use and evolution of ALOs.
8. The community trusts that the environment can evolve alongside its own needs and culture.

Additionally, the lifecycle of ALOs includes issues of ALO creation and maintenance.
ALO design and creation presents the challenges of:

Issue	**20-yr research problems**	**15-year research problems**	**5-year research questions**	**2-year research questions**
Enabling trust	Educating (meaning it's a done deal, and not flexible any more?) the community and stakeholders in the processes and tool Continuing research and development of processes and tools as technology evolves	Evaluating and refining the processes; Developing the tools for facilitating these processes Tools for transparency in ALO creation, testing and maintenance	Defining and formalizing processes for involving communities and stakeholders in the creation and maintenance of ALOs Identifying a set of tools that will facilitate these processes	Identifying the stakeholders, their roles, responsibilities and needs Identifying gaps in current processes and tools
ALO identification, instrumentation, creation, incorporation into infrastructure	Educating the community on research and development of tools and processes for ALO creation Educating the community and stakeholders on the models and tools for formative, summative, and ongoing evaluation of ALOs Continuing research and development	Developing ALO creation and collaboration tools Developing tools and processes for testing, quality assurance, and stakeholder acceptance Develop operational models and tools for formative evaluations of ALOs	Developing process models of the creation of the different types of ALOs Identifying tools for collaboration, creation, and modifications. This includes authoring tools. Developing approaches to formally testing large-scale, distributed knowledge-based systems Develop prototype	Analyzing current practices and tools to identify strengths and shortcomings Synthesizing the current state-of-art and research on ALO creation processes and tools (e.g. CTA methodologies, ITS authoring tools). Develop a roadmap for authoring tool research Mapping out the space of different types or models of ALOs Identifying the stakeholders in ALO

			models and tools for formative evaluations of ALO	creation of the different types of ALOs and their roles and responsibilities Identifying ALO evaluation criteria
ALO maintenance and evolution	Educate the stakeholders of processes and tools for maintaining ALOs	Develop operational tools.	Develop ALO architecture that mitigates the effects of platform dependency and upgrades Develop ALO editing/repurposing tools. This would be a little bit different or perhaps derivative of authoring tool research, but we expect that the tool requirements for ALO modification, updates is different from the requirements for ALO creation. It therefore would be important to pay special attention to this problem as distinct from ALO creation tools.	Identifying the technological challenges of continued ALO use (e.g. platform independence, platform upgrades, forward compatibility with new technologies and platforms) and mapping emerging trends Identifying the technological challenges of modifying/evolving ALOs as the needs, knowledge, and cultural perspective of the community changes or transfer of ALOs across communities
LMS infrastructure (student profiles, links between ALOs, feedback and adaptivity mechanisms)				

- Knowledge engineering/cognitive task analysis (CTA): This is a recognized bottleneck in the development of knowledge intensive systems. Often the CTA process uncovers latent knowledge that the community was not aware of. This feeds back into the trust issue.
- Authoring: Different segments of the community must be able to repurpose an ALO to suit their particular needs and culture.

- Determining and controlling ownership of ALOs. Needs to be dealt with in ways that motivate the community to collaborate and build on each other's work.
- Freedom of use: People can choose to use or not use ALOs as suits them.
- Communities need to be included in the design of superhighway infrastructure, social networking capabilities, and the ALO allocations.
- Testing and quality assurance: The ALOs will play a significant role in the education well-being of the community. Testing and quality assurance is a very important consideration.

ALO maintenance presents the challenges of:

- Testing and quality assurance
- Evolving the ALOs as the communities' knowledge of the world evolves
- Updating the ALO as knowledge evolves, and detecting and resolving obsolescence
- ALO versioning and archiving

This table lays out a roadmap of Computer Science research problems arising out of the high-level challenges discussed in this section. We recognize that the lifecycle loop impinges heavily on systemic process, cultural effects, and requires participation, dialog and cross-fertilization across a number of research disciplines including social scientists and policy-makers. For the purpose of this document, we restrict the recommendations we make to the fields of computer science and the learning sciences, with the assumption that there will be parallel efforts to identify and fund research in the other connected areas.