Proceedings of the First Adaptive Instructional System (AIS) Standards Workshop

March 2018
Orlando, Florida

Edited by:
Robert Sottilare
Avron Barr
Arthur Graesser
Xiangen Hu
Robby Robson

Part of the Adaptive Tutoring Series

Proceedings of the 1st Adaptive Instructional System (AIS) Standards Workshop

Edited by:
Robert Sottilare
Avron Barr
Arthur Graesser
Xiangen Hu
Robby Robson

Printed in the United States of America

First Printing, September 2018

U.S. Army Research Laboratory

Human Research & Engineering Directorate

Orlando, Florida

International Standard Book Number: 978-0-9977257-3-5

Dedicated to current and future scientists and developers of adaptive instructional systems

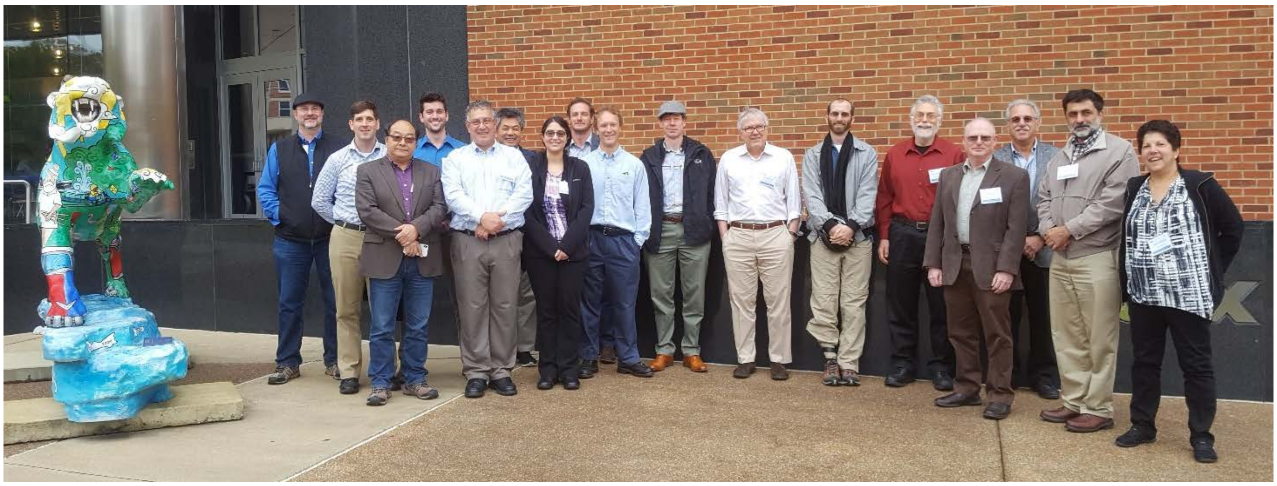
TABLE OF CONTENTS

From the Editors

Hi…

We thought it might be important to talk about how we got to where we are now in the adaptive instructional systems (AIS) standards process and expound on possible outcomes upfront in this proceedings.

In November of 2017, a group of scientists and technologists in the field of AISs met at the University of Memphis to discuss the efficacy of standards for AIS technologies that mainly included Intelligent Tutoring Systems (ITSs) and Intelligent Media. That group of about 15 participants (pictured below) included representatives from government, industry, and academia in the United States of America.



Pictured (L-R) at the University of Memphis (UofM) Fedex Center: Brandt Dargue (Boeing Company), Scotty Craig (Arizona State University), Xiangen Hu (University of Memphis), Drew Hampton (UofM), Bob Sottilare (ARL), Zhiqiang (Carl) Cai (UofM), Anne Sinatra (ARL), Ben Goldberg (ARL), Keith Brawner (ARL), Vincent Aleven (Carnegie Mellon University), Art Graesser (UofM), Andrew Olney (UofM), Eric Domeshek (Stottler-Henke Inc.), Robby Robson (Eduworks, Inc.), Avron Barr (IEEE), Peter Brusilovsky (University of Pittsburgh), and Jody Cockroft (UofM).

The meeting was organized by Professors Xiangen Hu and Arthur Graesser at the University of Memphis at the request of Dr. Robert Sottilare, lead for adaptive instructional sciences at the US Army Research Laboratory. The debate about AIS standards centered primarily on two presentations and a series moderated discussions:

- Dr. Sottilare set the stage and made a case for the benefit of AIS standards.
- Dr. Robby Robson and Avron Barr discussed the IEEE standards process, different types of standards
- Drs. Hu, Graesser, and Olney moderated discussions related to timelines and opportunities for AIS standards, and when standards are a good idea.

At the end of the meeting, there was a clear consensus to continue the investigation and Avron Barr, Chair of the IEEE Learning Technologies Steering Committee (LTSC) asked Dr. Sottilare to form an AIS standards study group under the LTSC. The following month (December 2017), the LTSC approved the formation of a six month AIS Standards Study Group to investigate the possible market need for standards across AISs. Several interactions with AIS stakeholder communities point to a broad interest in AIS standards. Discussions to date indicate opportunities to influence their affordability, usability, maintenance, interoperability and reuse, making them more appealing to the masses. In addition to the workshop documented in these proceedings, the AIS study group has approved workshops and sessions at the following AIS stakeholder events:

- AIS session at the 6th Annual Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium in Orlando, Florida, 9-11 May 2018
- AIS workshop at the Intelligent Tutoring Systems (ITS) Conference in Montreal, Canada, 11-15 June 2018
- AIS workshop at the Artificial Intelligence in Education (AIED) Conference in London, UK, 27-30 June 2018

The goal of these workshops/sessions are to expose a broader, international audience of AIS stakeholders to the potential of AIS standards and recruit them to participate in the IEEE process. AIS workshops are relevant to communities researching and developing solutions including users of broadly used AIS technologies including, but not limited to the Cognitive Tutor, AutoTutor, the Generalized Intelligent Framework for Tutoring (GIFT), Dragoon, ASPIRE-Tutor, and the Total Learning Architecture (TLA).

The proceedings of this workshop are intended to document the evolution of proposals for AIS standards. Nine papers were presented as part of six discussion session conducted over a period of two days. The talks, papers, and discussion are documented herein along with recommendations for next steps.

If you are interested in participating in the AIS standards activity, information about the IEEE P2247 Working Group, news and scheduled events are announced through the IEEE Adaptive Instructional Sciences (AIS) website at: http://sites.ieee.org/sagroups-2247-1/members/. Please sign up today.

In closing, Dr. Sottilare would like to acknowledge the hard work and dedication of our program committee in making this 1st AIS Standards Workshop possible. Thank you all:

- Avron Barr, Aldo Ventures, IEEE LTSC
- Benjamin Bell, Eduworks
- Elyse Burmester, Dignitas Technologies
- Noel Cal, Dignitas Technologies
- Jody Cockroft, University of Memphis
- Paula Durlach, US ARL
- Art Graesser, University of Memphis
- Xiangen Hu, University of Memphis
- Rodney Long, US ARL
- Robby Robson Eduworks

We also want to thank the workshop moderators, speakers, authors, and participants (pictured below) for their significant contributions to this continuing debate about AIS standards:

- Daniel Barber, University of Central Florida Institute for Simulation & Training (author)
- Avron Barr, Aldo Ventures, IEEE LTSC (moderator, speaker, participant)
- Benjamin Bell, Eduworks (moderator, speaker, participant)
- Mary Jean Blink, TutorGen (participant)
- Ryan Baker, University of Pennsylvania (author)
- Gautam Biswas, Vanderbilt University (author)
- Keith Brawner, US ARL (author, participant)
- Elyse Burmester, Dignitas Technologies (participant)
- Zhiqiang (Carl) Cai, University of Memphis (author, participant)
- Noel Cal, Dignitas Technologies (participant)
- Jody Cockroft, University of Memphis (participant)
- Chad Coleman, Columbia University (speaker, author, participant)
- Jeanine DeFalco, US ARL (moderator, speaker, author, participant)
- Paula Durlach, US ARL (moderator, speaker, author, participant)
- J. Dexter Fletcher, Institute for Defense Analyses (IDA) (participant)
- Jeremiah Folsom-Kovarik, SoarTech (participant)
- Benjamin Goldberg, US ARL (participant)
- Arthur Graesser, University of Memphis (moderator, author, participant)
- Drew Hampton, University of Memphis (participant)
- Michael Hoffman, Dignitas Technologies (participant)
- Robert Hoy, Northrop Grumman (participant)
- Xiangen Hu, University of Memphis (moderator, speaker, participant)
- Jong Kim, US ARL (speaker, author, participant)
- Elizabeth Lameier, University of Central Florida Institute for Simulation & Training (speaker, author, participant)
- R. Bowen Loftin, University of Missouri (participant)
- Rodney Long, US ARL (participant)
- Daniel McCoy, Knowledge Net Consulting (participant)
- Ramkumar Rajendran, Vanderbilt University (speaker, author, participant)
- Lauren Reinerman, University of Central Florida Institute for Simulation & Training (author)
- Robby Robson, Eduworks (keynote speaker, participant)
- Vasile Rus, University of Memphis (participant)
- Anne Sinatra, US ARL (participant)
- Amy Sommer, USMA (speaker, author, participant)
- Robert Sottilare, US ARL (speaker, author, participant)
- Randall Spain, NC State University (participant)
- Richard Vanderbilt, Alion Science & Technology (participant)

Again, many thanks to all who contributed…

Regards,

Bob, Avron, Art, Xiangen & Robby

Editors, 1st AIS Workshop Proceedings

SECTION 1: WORKSHOP OPENING REMARKS

Opening Address of the AIS Standards Workshop

Dr. Robert Sottilare

U.S. Army Research Laboratory

WORKSHOP GOALS & EXPECTED OUTCOMES

Dr. Sottilare welcomed the participants, thanked the workshop program committee for their contributions, and then discussed related workshops and sessions at the Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium in Orlando (9-11 May), the Intelligent Tutoring Systems (ITS) Conference in Montreal (11-15 June), and the AI in Education Conference in London (27-30 June). Next, he laid out a set of goals for the workshop:

1. Help us define what is included in the technology area known as Adaptive Instructional Systems (AISs) and what is not
2. Identify problems that AIS standards might help solve
3. Identify opportunities for AIS standards (near term, low hanging fruit)
4. Discuss the benefits/merits of candidate AIS standards so they can be ranked
5. Publish our process and our findings: Proceedings of the First Workshop on Adaptive Instructional System (AIS) Standards

Ground rules were also set for the moderators, presenters and participants. Discussion moderators were asked to provide leading questions to the workshop audience. The goal during discussions was to hear as many people as possible, but the moderators would determine when to move on to another topic. The paper presenters were asked to complete their presentations in 20 minutes to allow 5-10 minutes for focused discussion on the ideas presented and transition to the next speaker. The participants were asked to be concise and courteous.

Dr. Sottilare also told participants that they should feel free to disagree with any idea presented. He noted that if the discussion in any session should stall, we would go directly to the papers in that session and come back to the discussion later in the session. Two PC members were designated to take notes, and capture actions and decisions. Finally, the agenda for the two day workshop was reviewed with the workshop participants to see if there were any changes, comments, or questions.

Standardization vs. Innovation: Seeking Common Ground

Dr. Robby Robson

Eduworks

KEYNOTE ADDRESS SUMMARY

Dr. Robson's keynote address highlighted why standards are important and when they make sense to implement. He presented three motivations for standards: 1) to promote interoperability (physical, data, and semantic); 2) to develop markets and structuring to enable supply chains, to lower barriers to entry, and to define product categories; and 3) for customer protection to improve reliability, repeatability, and quality. Dr. Robson reviewed several types of standards:

- Data Models
  - Enable separate systems or components to provide services to each other
  - Define the data that can be exchanged, how it is formatted, and what it means
- Interface Standards
  - Enable separate systems or components to provide services to each other
  - Define how to establish, maintain, and use a connection to exchange data
- Reference Models
  - Enable people to understand the architecture of complex systems
  - Define the components (or layers), what each does, and how they fit together
- Process Standards
  - Improve the reliability of processes and the quality of their outputs
  - Define best practices, their requirements, inputs and outputs, and metrics
- Category Standards
  - Provide labels for categories of products and help ensure that customer expectations are met
  - Define category labels, the capabilities that each represents, and related conformance criteria

Dr. Robson noted that standards made sense when: 1) there is a market need or opportunity, 2) there is common ground, and 3) there are benefits to a shared solution. Standards should also protect intellectual property and any unique features of products that provide a vendor a competitive advantage. By doing this, competitors in the market are encouraged to implement the standard.

Dr. Robson cautioned the workshop participants that standardization can be harmful when: 1) there is no adoption, 2) there is adoption, but it stifles innovation or competition, or 3) there is adoption, but it promotes a bad solution. Reasons for lack of adoption include: no market need or industry participation in the development of the standard, no technical benefit, or the standard is too hard, expensive, or complex to implement. Standards may stifle innovation or competition when there is collusion among competitors or dominance by one provider, or when there are unknown patents restricting the process. Finally, standards might promote poor solutions when there is consensus among producers but not consumers (or vice versa), or there is inadequate or faulty technical input during the standards process. He noted that standards processes are designed to avoid these pitfalls.

Next, Dr. Robson presented some general thoughts on AIS standardization and shared some ideas on common ground in AIS standards with respect to data. He also presented some thoughts on where he believes the best opportunities for AIS standards might lie.

He stated that AISs are instructional systems first and adaptivity (and interactivity) is a spectrum. He asserted that standards should support arm's length operation and leave the inner workings alone since that is the secret sauce. Dr. Robson urged for a focus on the essential operations. For example, lots of data could be exchanged between AISs, but what must be exchanged and why should it be exchanged? We should be thinking about maintenance, implementation costs, and lifecycle costs of standards when considering the need for standards. He stressed that standards should improve reliability and not increase complexity.

In examining AIS data and looking for common ground among AISs, he posed several questions for the workshop and AIS standards study group to consider:

- What data are required by most learning systems?
  - Who is the learner?
  - What are their goals?
  - Where are they in a curriculum?
  - What are their preferences and accessibility criteria?
- What data are required by most adaptive systems?
  - What does the learner already know and what don't they know?
  - What can the learner already do and what can't they do?

Similarly, he examined AIS data looking for things outside of what he considered to be common ground and listed three areas for exclusion from the standards process: 1) micro-adaptive portions of learner models which are not needed for interoperability and often contain secret sauce, 2) pedagogical/expert models which represent the operation of the instructional inner loop which are not needed for interoperability and often contain secret sauce, and 3) super fined grained domain model (e.g., knowledge components) which he believes would be little consensus to standardize and often include secret sauce.

Finally, Dr. Robson discussed potential AIS candidates for standardization:

- Domain models – the public structure of a domain might be standardized along with models of competency
- Learner profiles – the structure of a learner profile in terms of learning goals (pathways), learning history (badges, outcomes, completions), learner knowledge and skills (competencies)
- Processes that are not learning specific
  - Identity management
  - Security
  - Privacy protection
- Areas where consumers are confused
  - Levels of interactivity (as defined in learning objects metadata - LOM)
  - Levels of adaptivity (as suggested by Brusilovsky)
  - Types of ITS and AIS (as suggested by Barr)
- Developers, acquirers, and maintainers need help
  - Reference models (see IEEE 1484.1-2003, open systems interconnection - OSI, Van Lehn)
  - Line Replaceable Units (LRUs, as suggested by Sottilare)
  - Common vocabularies (as suggested by several people)
- Performance and quality matter

    - Scalability and other performance metrics
    - User experience metrics
    - Pedagogical effectiveness

All that said, Dr. Robson voiced his picks for candidate AIS standards in terms of yes, maybe, no, or done elsewhere:

- Domain models (yes)
- Learner profiles and associated reporting (yes)
- AIS categories (yes)
- Reference models and LRUs (maybe)
- Component & service registration (technical - maybe)
- System profiles (beyond categories - maybe)
- Effectiveness ratings (no)
- Pedagogical models and expert models (no)
- Best practices (no)
- Results reporting (done elsewhere)
- Identity management (done elsewhere)
- Privacy and security (done elsewhere)
- Content management and resource metadata (done elsewhere)
- Competency frameworks, badges, credentials (done elsewhere)

DISCUSSION

Dr. Sottilare noted the importance of this keynote presentation and thanked Dr. Robson for agreeing to provide it. Dr. Robson's background and experience in the IEEE standards process is invaluable to the AIS community.

Dr. Sottilare largely concurred with Dr. Robson's recommendations, but varied slightly on what might be pursued as an AIS standard first. He asserted that any areas where consumers are confused might serve as an indicator of a standardization need and suggests that an AIS reference model might be the most appropriate place to start.

Section 2: Discussion on Adaptive Instructional System (AIS) Product Categories

AIS Product Categories in 2020

Mr. Avron Barr

Aldo Ventures, Inc. & IEEE Learning Technologies Steering Committee

PRESENTATION SUMMARY

Mr. Avron Barr provided a lead-in presentation in support of our discussion on AIS product categories. Mr. Barr described the IEEE Standards Association (SA) principles as noted below:

- Due process: IEEE asks us to follow highly visible procedures which are set at the IEEE SA level, the sponsor level, and the work group level; IEEE also asks that the process be transparent.
- Openness: All interested parties can actively participate in the standards process
- Consensus: A clearly defined percentage is required for approval (e.g., 75%)
- Balance: All interested parties are represented in the process and that no single party has an overwhelming influence
- Right of appeal: Anyone can appeal any decision at any point

Mr. Barr also reviewed the IEEE SA's process and you can learn more about finding and developing IEEE standards at: https://standards.ieee.org/develop/govern.html. He also stressed the need for patience in the standards development process and cautioned participants that publishing a standard is just the beginning of the process noting that only the market can truly implement standards as their value is recognized in the resulting products in the marketplace.

His final thoughts on the process were stern, but based on experience. He provided the following advice to the workshop participants:

- Don't start a project if you are not solving a market problem: proprietary, monolithic, and single-vendor solutions don't need standards; this is not vanity publishing… it's hard and any people will put in many hours to realize a standard
- Much of the work is in making sure the document is complete, accurate, and unambiguous in the eyes of a broad spectrum of stakeholders
- Standards documents are important to both product developers and policy makers, but adopters and their IT staff and system integrators just want things to work (e.g., clarity, easy implementation, test suites, product certification).
- If you want to learn more, IEEE has resources:
    - IEEE Standards University: http://www.standardsuniversity.com
    - IEEE Standards Lifecycle: http://standards.ieee.org/develop/index.html

Next Mr. Barr spoke about AIS product categories and posed the following questions to the workshop participants:

- What are the current AIS product categories?
- How do we define their boundaries?

- Must AISs actually teach, or can they help learners and teachers in other ways?
- Are there different types of adaptivity that might differentiate product categories in the future?
- Is there some measure of "adaptive power"?

Some examples of AIS products include:

- Massive, monolithic solutions: Carnegie Learning, ALEKS, …
- Point solutions: Alelo's cultural tutors
- Personal Assistants and Recommenders: Mari, PERvasive Learning System (PERLS)
- Embedded AI in simulations (e.g., virtual reality environments and games)
- Platforms for smart learning activities: Open edX, Carnegie Learning, (GIFT), (ADL's TLA)
- Adaptive testing, assessment

In discussing AIS Product Evolution, Mr. Barr asked participants to project what types of products might come to market in the next 2-5 years?

- Separate ITS components for sale
    - Common student model elements kept in a personal app
    - Student data locker
    - Smart learning activities that plug into a platform
    - Plug-in pedagogies, domain models
- Tools for authors and publishers
- Tools for installation managers and teachers
- Combination products: Mari, for example, is a student model, a activity referatory, and a personal assistant

He also asked "What new kinds of products are we likely to see?"

- New technologies: affect recognition, robotics, IoT, immersive, blockchain, 5G, …
- Changes in education and training

With respect to how AIS might be categorized, Mr. Barr asked "How will the anticipated evolution of related product categories impact the types of standards needed?"

- Learning activities: immersive, online, mobile
- Learning Record Store
- Student data locker
- Robo-graders
- Repositories/aggregators for smart learning activities
- Competency management, evidence of mastery
- Analytics engines (dashboard, early warning, …)
- Learning management system, student info systems, …
- Personal tutors and AI assistants: Alexa
- Management/scheduling apps for teachers, students, parents, ...
- Identity and data security

As way of comparison, he elaborated on the different types of learning technology markets that exist today:

- Enterprise training
- Education
- K12
- Higher Ed
- Job and trade prep
- Professional certification
- Individual, life-long learning

DISCUSSION

Mr. Barr reiterated that the goal of the standardization process is to ensure everyone is heard and there is a consensus – anyone can participate and appeal. The IEEE is *constructed*, much like a jury; but only the MARKET can establish a standard! A leading question was then provided to the participants to motivate discussion: What is the key to defining what an AIS is (and isn't)?

- Dr. Fletcher stated that *individualization* is key because AISs adapt to individual learners
- Dr. Graesser stated grain size is an important consideration; at what level is the system adaptive to (performance, progress toward objectives, etc.)?
- Dr. Sottilare noted that adaptive models need to be more intelligent about the learner (ie. Affective states)
- Dr. Brawner questioned the need for minimally adaptive systems versus maximally adaptive? Instructional component should be a key consideration – it should DO something
- Dr. Fletcher stated that categories of competencies should drive AIS standardization
- Dr. Bell stated that the focus on what you think makes an AIS different from other instructional systems
- Drs. Sottilare/Graesser noted that the goal is always competency, if a production rule includes affective state measures then that will determine how the student should be instructed – should be selected by experience
- Dr. Sottilare stated that AIS needs to reinforce its own decisions; system should be able to recast model over time – if it becomes too domain specific, it becomes too hard to standardize; domain STRUCTURE might be the key
- Dr. Bell stated that AISs need to be able to conform to functionality
- Dr. Sottilare stated that AIS standards should provide opportunities to not only bring in open products but proprietary products as well; your system has a certain threshold, integrating should require a cost, if they don't want to pay they should conform to standards

Section 3: Discussion on Opportunities for Interaction Standards

Opportunities for Interaction Standards between Common Components, AISs and Users, and AISs and External Systems

Dr. Jeanine Defalco

U.S. Army Research Laboratory

DISCUSSION

- The group was encouraged to break into pairs to come up with a topic they feel is most important to standardizing interoperability and creativity; these included:
  - Cultural Interactions
  - Not reinventing the wheel while the wheel is spinning
  - Generalization
  - Learner Data and Content Metadata
  - Organizational Culture and Technology Integration
  - Individuals, Collections of Individuals, Teams
  - Competencies, Speed and Accuracy, Appropriateness
  - Interoperability
  - Leveraging Emergent Standards from Integration Efforts
  - Industry vs. Academia, Cultural Acceptance
  - Integration of Innovation to Answer Gap Questions
  - Data Aggregation
- Group Chose Top Three to discuss: Universal KSA's, Interoperability, and Reinventing the Wheel
- Universal KSA's
  - Dr. Graesser stated that when adapting to humans, we should start with a large set of KSA's and populate the matrices as the field progresses, with both domain dependent and independent
  - Dr. Goldberg noted AIS should include things you can track over experiences – interoperability through some set of core states, maybe emotions for example
  - Rus identified the need to specify what input is needed in order to produce the desired outcome – what should we provide as input that will generate the learner model characteristics? It should go to a finer grain than just competencies, possibly plug and play architecture? Suggested using a *cognitive task analysis* that should specify what inputs are required to produce desired learning outcomes
  - Dr. Sottilare asked "What do I need to know about the learner to make effective pedagogical decisions? A learners attitudes, behavior and cognition are not known upfront; referenced work by Dr. Robert Neilson; important to determine whether an individual is ready to learn or has learned
- Interoperability/Reinventing the Wheel
  - General discussion focused on gaining understanding of standards relating to other standards/standards building on standards
  - Mr. Barr suggested competencies might be a critical standard set to begin looking into
  - Dr. Durlach suggested looking into metadata standards, to which Mr. Barr identified Learning Object Metadata Standards as part of this area to look into; Dr. Hu has

experience using LOMS and integrating new components such as what kind of learner activities the user wants to see in the system

- No apparent existing LMS standard could be identified through discussion; various disciplines (DoD, academia, industry) produce different kinds of data, authoring interfaces, etc. , Mr. Barr identified SCORM and mentioned the possibility of other standards regarding LMS's that maybe no one else is using; cultural differences should be strongly considered for implementation

Exploring Standardization Opportunities by Examining Interaction between Common Adaptive Instructional System Components

Dr. Robert A. Sottilare[1] and Dr. Keith W. Brawner[1]

US Army Research Laboratory[1]

Topical Area: AIS Design and Interaction

INTRODUCTION

Most instructional systems today are non-adaptive or low adaptive, not tailored to optimize individual or team learning. In computer-based training (CBT), which is low adaptation, most of the tailoring in the instruction is based on learner knowledge. There are some objectives and measures to assess how a learner progresses toward those learning goals. Military organizations often distill these objectives down to *tasks*, *conditions*, and *standards* where the task is the experience or what is to be learned (e.g., how to read a map), the conditions represent the context under which the instruction takes place (e.g., outdoors in varying degrees of ambient light), and the standards are the measures of success (e.g., ability to identify features or elevations on a two dimensional paper map). The adaptation is low because the CBT reacts only to errors made by the learner. Other than a model of the task, the CBT has no understanding of the learner's states (e.g., engagement level, emotional state, interest and preferences) or their effect on his learning.

As part of background information provided in this paper, we propose a notional definition for adaptive instructional systems (AISs): *computer-based systems that guide learning experiences by tailoring instruction and recommendations based on the goals, needs, and preferences of each learner in the context of domain learning objectives*. Computer-based instructional strategies may take many forms, but generally center around two adaptations: 1) to guide the level of support (scaffolding) or 2) changes to the difficulty level of the content.

Per Vygotsky (1978), the zone of proximal development (ZPD) examines the difference between what a learner can do without support, what the learner might do with support and what he cannot do even with significant support. Vygotsky also stresses compatibility between the learner's domain competency and the difficulty level of the content presented and discussed in that domain. Learners who experience content well below their domain competence are often bored, a negative learning state. While learners who experience content well above their domain competence are often anxious, a negative learning state. Ideally, we would want to avoid these states, but if they occur, there are opportunities to transition to new positive learning states through support or adaptations to the content level of difficulty.

In Intelligent Tutoring Systems (ITSs) this is accomplished by acquiring behavioral or physiological data from the learner to assess or project recent learner states. These states can be used to guide ITS decisions on the selection of domain-independent strategies (e.g., support, prompts, and reflective dialogue) and ultimately selection of tactics or specific domain-dependent actions by the tutor. This sequence of interactions resulting in optimal tutoring decisions has been described in the *learning effect model (LEM)* for both individual learners and teams of learners (Sottilare, 2012; Fletcher & Sottilare, 2013; Sottilare, 2013; Sottilare, Ragusa, Hoffman & Goldberg, 2013; Sottilare, et al., 2017). We are suggesting that the interactions between common ITS components, which are often arranged as encoded messages, might be our best near term opportunity for standardizing AISs. Interactions within and external to the ITS involve its four common components: learner models, instructional models, domain models and a user interaction model.

The learner model functions to acquire and interpret data sourced by the learner either through self-report mechanisms, sensors (behavioral or physiological) or historical records from learner record stores (LRSs) or learning management systems (LMSs). Instructional models vary in construct, but generally provide mechanisms to assess progress toward learning objectives and then provide appropriate recommendations with respect to the next experience or task for the learner. Domain models define the experiences, objectives, measures, assessments, and interventions (e.g., feedback, dialogue) of the domain under instruction along with learner attributes (e.g., motivation, grit, prior knowledge, assessed knowledge or skill) which form the basis to adapt the instruction. The interventions by the tutor and the learner are managed through a user interface. In tutorial frameworks like the Generalized Intelligent Framework for Tutoring (GIFT; Sottilare, Brawner, Sinatra & Johnston, 2017; Sottilare, Brawner, Goldberg & Holden, 2012), the tutor-user interface (TUI) manages communications between the tutor and the learner(s), but also serves to deliver various forms of content (visual, aural, textual, and hybrid content).

PROBLEMS, OPPORTUNITIES, AND DECISIONS

In this section, we have identified problems, approaches, design decisions, and associated benefits/opportunities in developing standardized messages to represent the interactions between common AIS components including models of the learner, the instruction, the domain, and the user interface.

Problem: How to Promote Reuse between AISs

Santos and Jorge (2013, p.271) note: "*Because of interoperability issues, intelligent tutoring systems are difficult to deploy in current educational platforms without additional work. This limitation is significant because tutoring systems require considerable time and resources for their implementation. In addition, because these tutors have a high educational value, it is desirable that they could be shared, used by many stakeholders, and easily loaded onto different platforms.*"

At the heart of the problem is *interoperability* which Wegner (1996) defines as: "*the ability of two or more software components to cooperate despite differences in language, interface, and execution platform. It is a scalable form of reusability…*" Therefore, logic dictates that any enhancement to make systems more interoperable will likely result in higher reuse of components, lower development costs, and more collaboration in both the research and development of AISs. Such is the point of this workshop, and this paper seeks these benefits among the models of learners, instruction, and domains.

Opportunity: Standards for Interoperability and Reuse

A goal for AIS technology is to be affordable, effective and ubiquitous for a variety of instructional domains. Most tutoring systems are tied to a single instructional theory which limits their effectiveness in domains outside the one they were designed to support. This makes the prospect of building AISs for a broad array of instructional domains an expensive proposition. In recent years frameworks like GIFT have sought to provide a single authoring tool to construct AISs in a broad array of domains across cognitive, affective, psychomotor, and social tasks, but GIFT also has limitations within its authoring and instructional tools. The engine for managing adaptive pedagogy (eMAP), is currently based on Merrill's Component Display Theory (1983) with a planned expansion to Chi's ICAP Framework (2014) in 2018. It remains unclear whether this instructional engine will be robust enough to cover all the domains of instruction that designers would like to implement as AISs.

The problem is the time and cost to integrate shared AIS components or modify existing AIS components so they can be used in other AIS applications. The barrier to easily integrating AIS components is a full understanding of what information each component requires to make decisions and to identify a source for this information. Standardizing messages will facilitate the development of a standard set of AIS functions (e.g., learner model initialization) allowing the use of compliant system components in other AISs. For example, an instructional model in one AIS featuring Merrill's Component Display Theory as an instructional policy could be exchanged for another compliant model featuring Gagne's nine instructional events as an instruction policy or vice-versa assuming the information required by each instructional model is the same or at least of the same structure. This is not idle writing, as automated tools have up-converted EMAP-based GIFT courses to iCAP-based GIFT courses without breaking the underlying content, but providing new models of instruction for existing content at practically-no cost.

The core components of a tutoring systems are widely agreed upon. Murray's seminal review work from 1999 included the components of the student interface, domain model, teaching model, and student model (Murray, 1999). Woolf's review of ITSs divides the space into models of student knowledge, domain knowledge, tutoring knowledge and the communication of knowledge (Woolf, 2010). VanLehn's (2006) review describing the behavior of tutoring systems mentions "that although tutoring systems differ widely in their task domains, user interfaces, software structures, knowledge bases, etc., their behaviors are in fact quite similar". GIFT, as a system, includes a domain model, a pedagogical model, a learner model, and interface model, each of which map directly to these items, and whose design has been influenced by a series of expert workshops. The Army Research Laboratory has published a volume of recommendations on the subject of each of these modeling methods in an effort to capture the best practices from government, academia and industry. It seems clear that each ITS must, at a minimum, include a model of the domain of instruction, a model of the learner that it is instructing, and a model of instruction of what content to deliver to the learner.

The ability to exchange instructional models in an AIS allows for greater flexibility in using the same basic framework to provide instruction in a variety of domains, but without the cost of building each AIS from scratch. Even partial interoperability that builds over time would serve to reduce AIS development costs, but at what level should compliance be implemented for AIS standards?

Design Decision: Compliance at the Lowest Replaceable Unit (LRU)

A key decision that designers face in addressing opportunities for AIS standardization is at what level to implement compliance. Seeking to implement standards at the lowest functional level possible means more easily implemented standards, but also a great number of standards. This may or may not lead to value-added solutions or reduce the cost of integration of similar components (e.g. one instructional model or another instructional model). On the other hand, implementing standards at a much higher (generalizable) level can also be difficult since the standard must cover a higher number of use cases. Looking at this decision more closely (Figure 1), we see that the lowest replaceable unit (LRU) is at the component level (learner, instructional, and domain models). Regardless of what processes may be running within the model, the inputs and outputs remain the same. If we replace a component, the assumption is that the processes within the component will be able to use the inputs to generate appropriate outputs.

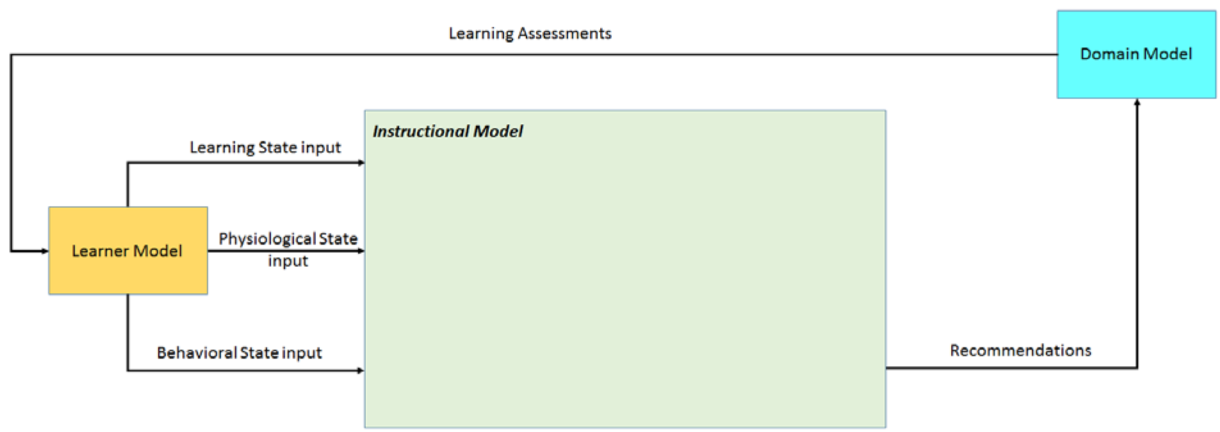


Figure 1. Example of Component Level Interoperability

Changing the level of the LRU to the process level (Figure 2) means each process must be compliant with the both the component level inputs/outputs and the process level outputs which create additional constraints when considering the appropriate level of interoperability.

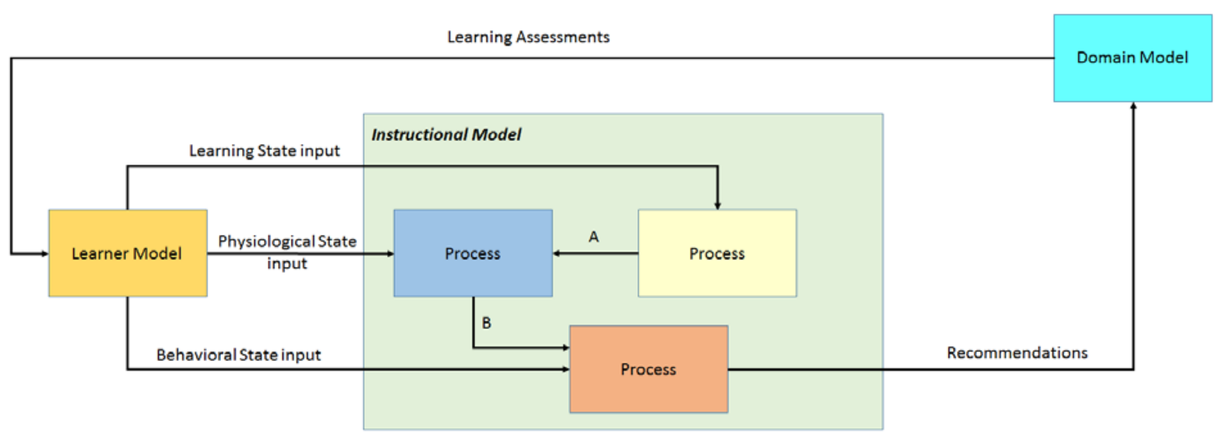


Figure 2. Example of Process Level Interoperability

Problem: How to Compare AIS Models, Components, and Processes Objectively

Currently it is very difficult to compare the efficiency or effectiveness of instructional systems that teach the same content. Sottilare and Ososky (2017, p.237) note that the "difficulty in authoring AISs vary with the complexity of their objectives, content, measures, assessments, and instructional decisions". Just as in comparing the structure, complexity is a factor. We argue that the comparison of systems, components, and processes within AISs would be facilitated by the use of standard message sets to relay information between components in AISs.

This lack of interchange hinders the field in two primary manners. Scientifically, the field is hindered by not being able to fairly compare models of instruction, content, and student modeling. Using a model of pedagogy as an example, AutoTutor uses natural language dialogue to didactically instruct students and ask questions in order to get students to think deeply about problems (Graesser et al., 2012). The Cognitive Tutor is guided by ACT-R theory and observed learning curve data in order to make optimal next task selection and feedback (Ritter, Anderson, Koedinger, & Corbett, 2007). What would the effect be if the Cognitive Tutor were to ask introspective questions in the manner of AutoTutor instead of selecting problems? What would be the effect of problem-based dialogue in AutoTutor be? Would a Cognitive Tutor probabilistic model of concepts aid AutoTutor in question selection? Would AutoTutor see learning gains within the math domain? Could Cognitive Tutor be effectively applied to conceptual learning of middle school science? Questions such as these cannot be simply answered through the interchange of models – they require the creation of a whole system. It is not possible to simply swap the models, or to opt to switch to a commercial competitor.

Opportunity: Standard Messages to Aid Comparison of Component Performance

Standardizing messages will potentially facilitate growth, development, and evaluation of the effect of various elements of AISs on learning outcomes (e.g., performance, retention, and transfer of skills). In other words, it will allow a standard of measure for comparison. As noted earlier in this paper, the instructional model provides both recommendations for next steps or experiences and at least in the case of GIFT, a domain-independent strategy that constrains the selection of actions by the domain model. Comparing instructional model performance is difficult because inputs from the learner model of two different AISs may differ and therefore the strategy recommendations for the same learner under the same conditions of instruction might also vary (i.e., under learner state {novice}, one model gives a hint, but another assigns remedial content). It may also be that one system provides recommendations while another may not (i.e., under learner state {expert}, one model gives a hint, but another gives nothing and instead waits for a summative review to provide feedback). By standardizing messages to represent component interactions, it allows investigators to examine how comparable components operate under an identical set of conditions. AIS designers may find that as with many models their boundary conditions (area where the models perform best) point to the use of multiple models to cover the diversity of instructional domains. As an example, for a learner in this state, is it best to provide a hint, provide remedial content, or wait until a later summative review?

Design Decision: Where to Start with Standardization of Messages

We recommend an analysis of the information required by AIS common components to begin formulating an ontology as a basis for selecting candidates for standard messages. Based on experience with GIFT and in reviewing the literature, Table 1 lists some of the most frequently recurring examples, but is intended to be illustrative, not exhaustive. Note that within the instructional model, many of the inputs and outputs could be optional, but a system with fewer instructional outputs is less likely to be interactive (e.g., page turner) according to multi-media instruction models (Schwier & Misanchuk, 1993; Mayer, 2005) and certainly less adaptive. Whether you agree or disagree, this table provides a starting place for component-level standardization discussion, and invites comment.

Table 1. Sample of Interaction between AIS Models

- Domain Model
  - Input
    - Requests for action (from Instructional Model)
    - Feedback associated with concepts
    - A model of domain tasks, conditions, and standards (measures)
  - Output
    - Learner assessments (to Learner Model)
- Learner Model
  - Input
    - Learner assessments for each learning objective or concept (from Domain Model)
    - Learner State representation (from Domain Model or derived from data)
    - Sensor data (if applicable)
    - Longer term data (if applicable)
- Instructional Model
  - Input
    - Learning State representation (from Learner Model)
      - Cognitive state of the learner
      - Performance expectations (above, below, at) for each concept
      - Predicted future performance based on competency model
    - Physiological State representation (from Learner Model)
      - Derived emotional, physical states (e.g., fatigue)
      - Physiological stressors
    - Behavioral State representation (from Learner Model)
      - Derived attitudes or psychomotor performance based on primitive behaviors
    - Longer term learner attributes (from Learner Model or Learner Record Store)
      - Demographics and traits
      - Historical performance (competency)
  - Output
    - Request for course direction (to Domain Model)
    - Request for feedback (to Domain Model)
    - Request for scenario adaption (to Domain Model)
    - Request for assessment (to Domain Model)

DISCUSSION

Additional benefits of standardization of AIS messages might also include extensibility and flexibility. A standard set of messages could easily grow along with the development of new AIS functional capabilities, but still allow “old” and “new” components to be interoperable at varying levels. The opportunity for reuse will both reduce the cost of authoring AISs and encourage the sharing of ideas and AIS components. This allows system developers to be able to determine when and to what degree of interoperability they need to function within a given community.

Standard messages might also be triggers to call for services within and external to AIS components. Standard messages with associated arguments might also be used to trigger new processes. While it may be difficult to standardize AIS components, the processes running within those components could remain interoperable with other AISs based on standard messages and adjustable arguments allowing for small adaptations when sharing components across different AISs.

When a learner starts a tutoring session, the learner, instructional, and domain models must be initialized along with any sensors or gateways required to collect learner behaviors or allow interaction with third party software. The initialization process by itself might be a significant leap forward by allowing developers to share common data across various learner populations, instructional domains, and hardware/software platforms, but standard interaction between models and interfaces on a regular basis (event or time triggers) might also be useful in updating these models based on the learner's progress toward objectives in a current tutoring session or their achievements across various domain sessions with different learning platforms (e.g., GIFT, AutoTutor, Cognitive Tutor, PERLS). Outer loop recommendations (next steps within the session or what to do after completion of the current session) might also benefit from some standard messaging to be consumed by other systems in the distributed learning landscape.

It might be possible to develop some generic messages with enumerations to identify the specific purpose of the message. Similar to messaging in GIFT, sample generic messages are shown in Table 2.

Table 2. Sample of Generic AIS Messages

- Start/end domain session
- Discover module (learner, sensor, domain, gateway)
- Request course state
- Request action (hint, prompt, reflective dialogue, select question, deliver content)
- Read from Learner Model
- Write to Learner Model

CONCLUSIONS AND RECOMMENDATIONS

We are advocating for the development of an interaction ontology for AISs to guide the selection and structure of standardized messages representing interaction between the four common AIS models (learner, instruction, domain, and user-interface). Tables 1 and 2 provide a starting point for debate on which messages might be most useful and at which level of abstraction they should be implemented within AISs.

Based on experience, the creation of holistic tutoring system relies upon the creation of its parts – which has mostly been agreed upon in the literature. While we recommend the use of the GIFT Interface Control Document as a starting place in developing AIS standards, we also highly recommend designers review VanLehn's system actions to understand a broader representation of AIS technologies.

REFERENCES

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning out- comes. Educational Psychologist, 49(4), 219-243.

Graesser, A. C., D'Mello, S., Hu, X., Cai, Z., Olney, A., & Morgan, B. (2012). AutoTutor. In Applied natural language processing: Identification, investigation and resolution (pp. 169-187). IGI Global.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press.

Merrill, M. D. (1983). Component display theory. Instructional-design theories and models: An overview of their current status, 1, 282-333.

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. International Journal of Artificial Intelligence in Education (IJAIED), 10, 98-129.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. Psychonomic bulletin & review, 14(2), 249-255.

Santos, G. S., & Jorge, J. (2013). Interoperable intelligent tutoring systems as open educational resources. IEEE Transactions on Learning Technologies, 6(3), 271-282.

Schwier, R., & Misanchuk, E. R. (1993). Interactive multimedia instruction. Educational Technology.

Sottilare, R.A., Brawner, K.W., Goldberg, B.S. & Holden, H.K. (2012). The Generalized Intelligent Framework for Tutoring (GIFT). Concept paper released as part of GIFT software documentation. Orlando, FL: U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED). Retrieved from: https://gifttutoring.org/attachments/152/GIFTDescription_0.pdf

Sottilare, R., Brawner, K., Sinatra, A. & Johnston, J. (2017). An Updated Concept for a Generalized Intelligent Framework for Tutoring (GIFT). Orlando, FL: *US Army Research Laboratory*. May 2017. DOI: 10.13140/RG.2.2.12941.54244.

Sottilare, R.A. & Ososky, S. (2017, July). Defining Complexity in the Authoring Process for Adaptive Instruction. In *Foundations of Augmented Cognition* (pp. 237-249). Springer International Publishing.

Vanlehn, K. (2006). The behavior of tutoring systems. International journal of artificial intelligence in education, 16(3), 227-265.

Wegner, P. (1996). Interoperability. ACM Computing Surveys (CSUR), 28(1), 285-287.

Woolf, B. P. (2010). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Morgan Kaufmann.

ABOUT THE AUTHORS

Dr. Robert Sottilare *leads adaptive training research within ARL's Learning in Intelligent Tutoring Environments (LITE) Lab and is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT).*

Dr. Keith Brawner *is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT), and has been involved in ITS research since GIFT's beginning in 2011.*

DISCUSSION

- Drs. Sottilare and Brawner are essentially suggesting a plug and play idea, same concept as an LRU described in Dr. Robson's keynote; entire units should be able to be taken out and put into another system and remain operational
- Dr. Goldberg pointed out the possibility for error during translation within the standard
- Discussion identified the need for sensor standardization
- Mr. Barr pointed out that sometimes a standard can ultimately define a supply chain – something to consider as standards are crafted.

SECTION 4: DISCUSSION ON OPPORTUNITIES FOR LEARNER & TEAM MODELING STANDARDS

Opportunities for Individual Learner and Team Modeling Standards

Dr. Benjamin Bell

Eduworks, Inc.

DISCUSSION

- **Leading Question: How would a learner modeling standard help?**
  - Discussion focused on the need to consider both the process of learning for each learner and individual learner characteristics; what a learner model is/needs is not easily generalizable; need to define a domain carefully and consider multiple levels of analysis to determine which characteristics might be able to be standardized
  - Dr. DeFalco pointed out that the learner model is an interactive combination of who they are, what the content is, and how they are engaging the content, the live learning experience we are trying to model (individual, content, methodology) is both static and fluid
  - Competency alignment across models was identified as an important factor to consider – transfer of skills and experiences between ITSs, might be session/domain/situation dependent, short term vs long term should be considered
  - Prof. Graesser identified a Delta Standard in academia that models whether the learner is improving, provides a response to intervention and feedback about the learner when they are not reaching competency; delta information can be extracted via time stamps – something to consider
  - Dr. Rus suggested hierarchical categories for learner states could be beneficial for learner modeling standards because lower level details could be proprietary and high level state details might be interoperable; could provide the opportunity for users to utilize multiple business avenues
  - Learner models need to be updated over time; if a decay model is provided, the tutor needs a level of domain competence, Sottilare suggested standardization of how this data is reported and used by the tutor; Barr suggested breaking down the different types of data in learner models to determine how it should be queried; should begin with individual models before attempting team models;
  - Dr. Sottilare suggested starting by focusing on which learner attributes provide the largest pay off; Dr. Durlach mentioned a study that identified the highest average effect size of learner attributes was found in students attitude toward learning and their confidence in their ability to master the subject – something to consider

Standardizing Modeling of User Behaviors: Which Behaviors Matter?

Dr. Ryan S. Baker[1] and Mr. Chad J. Coleman[2]

University of Pennsylvania[1], Teachers College Columbia University[2]

Topical Area: Learner Modeling

INTRODUCTION

2004 marked the publication of what were arguably the first automated models of disengaged behavior within an adaptive instructional system (AIS) – models that captured gaming the system (Baker, Corbett, & Koedinger, 2004), or a subcategory of gaming, hint abuse (Aleven, McLaren, Roll, & Koedinger, 2006). Since then there has been a great deal of work that aimed to measure a range of disengaged behaviors (see review in Baker & Rossi, 2013), as well as behaviors that reflect positive self-regulation during learning (Shih, Koedinger, & Scheines, 2011; Tscholl et al., 2016). This work has shown not only that these behaviors can be modelled, but also that the resultant models can predict differences in learners' outcomes, both in the short-term (Rowe, McQuiggan, Robison, 2009; Sabourin, J., Rowe, Mott, & Lester, 2011) and in the long-term (San Pedro et al., 2013; Pardos et al, 2014). Finally, many of these behaviors are amenable to intervention to reduce disengagement and increase the frequency of positive behaviors (e.g. Baker et al, 2006; Arroyo et al., 2007; Bouchet et al., 2016), creating the potential to improve student outcomes.

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

Despite all of this research, the set of disengaged and engaged behaviors being modelled varies system-to-system, and the choice of behaviors modelled in specific research projects remains ad-hoc and unsystematic. There have been no efforts that the authors are aware of to standardize what set of behaviors should be modelled. This is partly because the field is still determining which behaviors matter, and new behaviors of importance are likely to emerge, especially as the design of AISs change to accommodate technologies such as virtual and augmented reality, and as learners themselves change over time. Nonetheless, many behaviors have been identified and studied in a range of learning systems – for example, models of gaming the system have been developed in at least 7 AISs that the authors are aware of, and models of off-task behavior have been developed in at least 9 AISs that the authors are aware of.

Given all of the work to detect specific known behaviors such as these, it seems feasible to start moving as a community to setting up a taxonomy of behaviors that should generally be modelled. Doing so, and having a shared framework for disengaged and engaged behavior in adaptive instructional systems, will enable different instructional systems to communicate with each other in a consistent way about how often students engage in specific behaviors known to be important, such as – for instance – gaming the system.

In this presentation, we will summarize the literature on these behaviors; which can be modeled; which have evidence for their importance to learner outcomes; and how robust and general that evidence is. We will also present thoughts towards a framework for modeling behavior in adaptive instructional systems, which would encapsulate widely-studied behaviors and have sufficient flexibility to accommodate new discoveries about user behaviors that emerge as adaptive instructional systems change, and the learners who use them change as well.

To start, this framework may be as simple as a list of common behaviors and their manifestations, with a commitment to representing behaviors using this taxonomy when possible, and embedding the taxonomy in a standard. This standard can and should be designed in a fashion where new behaviors can be tagged and added easily within the system, either for a specific system or as an addition to the standard by general agreement across a broader community of stakeholders.

The framework could support the communication of information between AISs with reference to this taxonomy, perhaps as a probability or estimated proportion of time that a student engaged in each behavior of interest (or an indicator if the behavior is inapplicable within a specific system). If a framework were designed in this fashion, it could accommodate outputs from a range of models in an algorithm/approach-neutral fashion. Whether or not a model of a specific behavior was based on feature-engineered machine learning, deep learning, rational modelling/knowledge engineering, or some as-yet-unknown approach, it could provide probability estimates or estimated proportions of time that would be meaningful and useful to the next AIS system the learner would encounter. The models themselves would not need to be domain or system-general; true system-generality is as yet uncommon in this space, with only Paquette et al. (2015) gaming detector being validated to be general in this fashion. The detector itself has only been validated to work effectively for three AISs (Paquette et al., 2015). Having a framework like this would be a step towards integrating information about a learner across multiple AIS systems, towards getting a broader picture of the learner, including their propensities to disengage in certain types of learning activities (perhaps some learners disengage with didactic materials whereas others disengage with games), their propensities to disengage more for some topics than others (an indirect measure of interest), and their overall path towards long-term engagement or disengagement with the learning subject.

DISCUSSION

One of the largest arguments against creating a cross-system framework for disengaged and engaged behavior is that new behaviors may still be discovered in the years to come, particularly as the systems we study change. However, we would argue that several behaviors have been shown to matter in several con- texts, and are worth developing standards for. Furthermore, it is not clear when or even if new behaviors will stop being discovered; the designs of AIS systems may continue to change for decades as new interaction technologies are developed, and changes in design may drive ongoing changes in how learners engage and disengage. This argues for a flexible and expansible framework rather than rejecting an attempt to codify and build on what is today known.

A second argument is that a cross-system framework will not be used, because different systems will be unable to trust models developed by others, and will not want to use evidence from other systems. It is true that cross-AIS transfer of knowledge about students has not yet become a practice. However, many systems have models of these behaviors that are published and heavily validated. Furthermore, the long-term predictive power of these behaviors (e.g. San Pedro et al., 2013; Pardos et al., 2014) suggests that a student who is disengaged in 8th-grade mathematics is likely to remain disengaged in 9th-grade. So, too, a military trainee's disengagement is relatively unlikely to simply disappear, from one month to the next. As such, retaining information on a learner's previous-year disengagement is likely to give a newly-encountered AIS system a head start on adapting to that learner's needs.

The biggest benefit to creating a cross-system framework and standard for engaged and disengaged behavior is the potential, discussed above, for AIS systems to exchange information and build information over time on student behavior. This has not yet been seen in the published literature; creating a standard may create awareness of this possibility and open developers towards having AISs work together to support students in engaging more effectively with learning content.

CONCLUSIONS AND RECOMMENDATIONS

We propose to create a standard for representing disengaged and engaged behavior, at a high level, within adaptive instructional systems. This standard will represent behavior as a set of categories, with sub-categories as appropriate; AISs that detect these behaviors will represent the behaviors in terms of these standard categories. Doing so will enable AISs to communicate with other AISs in a consistent fashion about whether a student is demonstrating a behavior, even if the two systems themselves detect the behavior in very different fashions (or even if one of the systems is itself unable to detect the behavior).

REFERENCES

Aleven, V., Mclaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. International Journal of Artificial Intelligence in Education, 16(2), 101-128.

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., ... & Woolf, B. P. (2007). Repairing dis-engagement with non-invasive interventions. In Proceedings of the International Conference on Artificial Intelligence and Education, 195-202.

Baker, R.S., Corbett, A.T., Koedinger, K.R. (2004) Detecting Student Misuse of Intelligent Tutoring Systems. Proceedings of the 7th International Conference on Intelligent Tutoring Systems, 531-540.

Baker, R.S., Corbett, A. T., Koedinger, K. R., Evenson, S. E., Roll, I., Wagner, A. Z., Naim, M., Raspat, J., Baker,

D. J. & Beck, J. (2006). Adapting to when students game an intelligent tutoring system. In Proceedings of the 8th International Conference on Intelligent Tutoring Systems (pp. 392-401).

Baker, R.S., Rossi, L.M. (2013) Assessing the Disengaged Behavior of Learners. In Sottilare, R., Graesser, A., Hu, X., & Holden, H. (Eds.) Design Recommendations for Intelligent Tutoring Systems -- Volume 1 -- Learner Modeling. U.S. Army Research Lab, Orlando, FL, pp. 155-166.

Bouchet, F., Harley, J. M., & Azevedo, R. (2016). Can Adaptive Pedagogical Agents’ Prompting Strategies Improve Students’ Learning and Self-Regulation? In International Conference on Intelligent Tutoring Systems (pp. 368-374).

Cocea, M., Hershkovitz, A. & Baker, R. S. J. d. (2009). The impact of off-task and gaming behaviors on learning: Immediate or aggregate? In Proceedings of the 14th International Conference on Artificial Intelligence in Education (pp. 507-514).

Paquette, L., Baker, R.S., de Carvalho, A., Ocumpaugh, J. (2015) Cross-System Transfer of Machine Learned and Knowledge Engineered Models of Gaming the System. Proceedings of the 22nd International Conference on User Modeling, Adaptation, and Personalization, 183-194.

Pardos, Z.A., Baker, R.S., San Pedro, M.O.C.Z., Gowda, S.M., Gowda, S.M. (2014) Affective states and state tests: Investigating how affect and engagement during the school year predict end of year learning out-comes. Journal of Learning Analytics, 1 (1), 107-128.

Rowe, J., McQuiggan, S., Robison, J. & Lester, J. (2009). Off-task behavior in narrative-centered learning environments. In Proceedings of the 14th International Conference on Artificial Intelligence and Education (pp. 99-106).

Sabourin, J., Rowe, J., Mott, B. & Lester, J. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In Proceedings of the 15th International Conference on Artificial Intelligence in Education (pp. 534-536).

San Pedro, M.O.Z., Baker, R.S.J.d., Bowers, A.J., Heffernan, N.T. (2013) Predicting College Enrollment from Stu- dent Interaction with an Intelligent Tutoring System in Middle School. Proceedings of the 6th International Conference on Educational Data Mining, 177-184.

Shih, B., Koedinger, K. R., & Scheines, R. (2011). A response time model for bottom-out hints as worked examples. Handbook of educational data mining, 201-212.

Tscholl, M., Biswas, G., Goldberg, B. S., & Sottilare, R. A. (2016). Detecting Metacognitive Strategies through Performance Analyses in Open-Ended Learning Environment. In Proceedings of the 24th International Conference on Computers in Education.

ABOUT THE AUTHORS

Dr. Ryan S. Baker *is Associate Professor at the University of Pennsylvania and Director of the Penn Center for Learning Analytics. He was the founding President of the International Educational Data Mining Society, has taught four MOOC instances, and was the co-developer of the Baker Rodrigo Ocumpaugh Monitoring Protocol. Baker has co-authored published papers with over 250 colleagues.*

Chad Coleman *is a doctoral student at Teachers College Columbia University and a participating member of the Penn Center for Learning Analytics at the University of Pennsylvania. His research interests center around the inter-section of human computer interaction, instructional systems, and learning analytics.*

DISCUSSION COMMENTS

- The authors suggested a triples format for a standard framework that represents behavior as a set of (mostly pre-defined) categories, defines the context of the behavior, and indicates the prevalence of the behavior – focus is on how the learner is interacting with the AIS, not how the AIS is engaged with the learner
- Research to follow that will introduce intervention but this is not currently touched on
- Discussion on what information should be required in the learner model provided varying opinions; Mr. Barr suggested the context of the data in this model is important to consider; Dr. Brawner disagreed and explained that what is required in a learner model does not necessarily have to have pedagogy on top of it; Dr. Graesser pointed out that if the student is not engaged in the material, they will not learn and mentioned that this system should detect this; Barr argued that this is a different type of standard; Dr. Goldberg suggested intervention at the instructor level; Dr. Fletcher pointed out a transferability issue between systems if the systems don't know context of data.

Standardizing Learner Modeling for Complex Domains Using Multi-Level Learner Modelling Schemes

Dr. Gautam Biswas and Dr. Ramkumar Rajendran

Institute for Software Integrated Systems, Department of Electrical Engineering and Computer Science, Vanderbilt University, 1025 16th Avenue South, Nashville, TN 37212, U.S.A.

Topical Area: Learner Modeling

INTRODUCTION

Complex decision-making tasks require learners to develop situational awareness, critical thinking skills and strategies, and a better understanding of how to evaluate potential solution approaches when working in problem spaces that are large, and often difficult to interpret. In addition, they need to develop abilities to assess and revise their evolving solutions to determine if they are making progress toward their learning and problem solving goals (Brehmer, 1992). In general, learners in such environments are made aware of the importance of setting goals, decomposing complex goals into sub-goals, and then given opportunities to devise strategies and apply appropriate cognitive skills to achieve these goals and sub-goals in the problem-solving environment (Wood & Bandura, 1989).

Intelligent learning environments (ILEs) can now support learning and training in complex problem-solving tasks in different domains. The cognitive and learning sciences literature emphasizes the need to develop cognitive skills, strategies, and metacognitive processes to become effective learners and problem solvers. However, novice learners often face difficulties when working on complex tasks, which impedes their progress. They may need adaptive scaffolding and feedback to overcome their difficulties and succeed in completing their tasks. These feedback mechanisms, and the learner models they are based on, have to consider the learner's performance as well as their learning behaviors to promote deeper learning and learning that may transfer across problems and domains.

We propose a connected, multi-level hierarchical learner-modeling scheme as a standard for modeling learners' cognitive skills, strategies, and metacognition along with their performance on assigned tasks. At the lowest level of this hierarchy are observable actions the learners can perform in the environment. For complex problems and decision-making tasks, a learner's proficiency must extend beyond individual cognitive skills to include (1) cognitive strategies, i.e., how to combine individual skills to achieve sub-goals and goals, and (2) metacognitive processes that are directed toward awareness, monitoring, self-evaluation, and reflection to guide the learning and problem-solving tasks. Such an approach allows for much richer analysis of learners' performance and behaviors in an ILE, and provides a framework for scaffolding learners at the skills, strategies, and metacognitive levels.

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

Complex problem solving and decision-making are essential skills to support human planning and decision-making processes. This is also accelerating the need for computer-based training environments that supplement traditional instruction for complex decision-making. Such learning environments are typically open- ended, i.e., they allow learners to explore different solution approaches and paths to solutions, while also helping them develop the abilities to evaluate and compare approaches, and trading off the pros and cons of different approaches before converging on the "best" or "most effective" solution. In general, open- ended learning environments (OELEs) focus on developing learners' cognitive skills, cognitive strategies and metacognitive processes that go beyond the acquisition of domain-specific knowledge (Biswas et al., 2016; Hannafin et al., 1994).

However, these environments can make high cognitive demands on learners, while promoting the development of strategies and metacognitive processes that enable learners to develop abilities to generalize and transfer their solution approaches across problems and even different domains. Less experienced learners often face difficulties because of the open-ended nature of the complex decision-making tasks. As a result, they may need scaffolding and support to make them aware of the requisite skills and strategies they need to develop to become proficient in their problem solving and decision-making tasks. In order to provide adaptive feedback, the OELE needs to

1. Track and record learners' progress in applying appropriate cognitive skills and strategies as learners work in the environment; and

2. Provide adaptive feedback to help the learners' overcome their difficulties in applying relevant skills, strategies, and metacognitive processes and become more proficient problem solvers in the domain.

Existing learner models in ILEs are mainly focused on learners' overall progress on applying domain concepts and performance in applying discrete cognitive skills. Most approaches do not consider how they combine these skills to form cognitive strategies to address the more complex aspects of problem solving tasks, such as goal setting, planning, and monitoring progress. In addition, for standardization purposes, learner-modeling need to consider components that support metacognitive tutoring. We believe that understanding learners' metacognitive processes require developing an integrated understanding of their awareness that leads to appropriate strategy selection, and their ability to monitor their own progress. This helps them review how well they are performing and then to refine and revise their problem solving approach as needed.

In order to measure learners' proficiencies on cognitive strategies that are derived from learners' proficiencies on cognitive skills and their relevance to the sub-goals and tasks they are currently working on, we have developed a multi-level hierarchical learner model with accompanying algorithms to track learners' progress in cognitive skills, cognitive strategies, and metacognitive processes (Rajendran, et al., 2017). In conjunction, we have developed instructional strategies that support adaptive scaffolding.

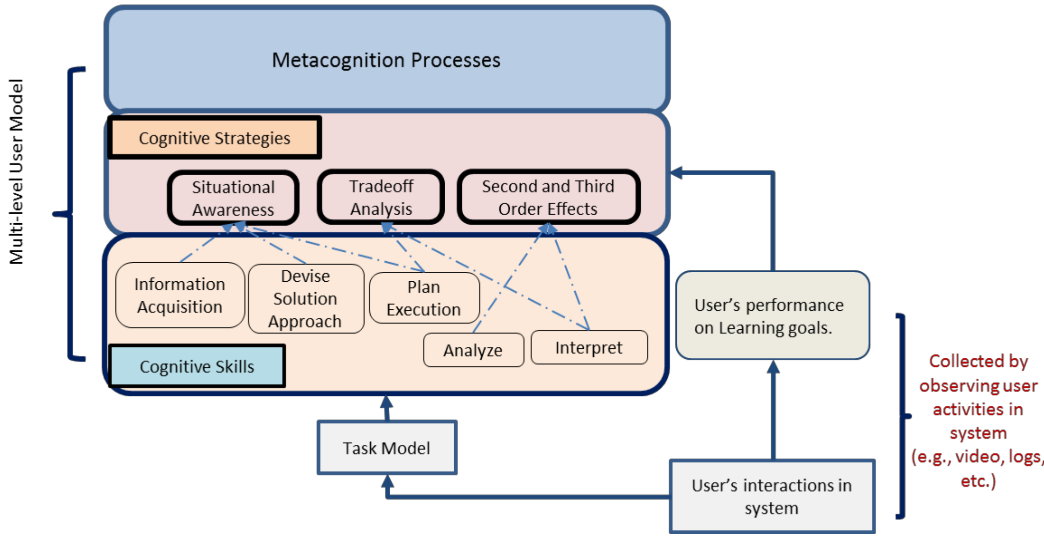


Figure 1. Three-Tier Hierarchical Learner Model (implemented in GIFT for training in UrbanSim)

Figure 1 illustrates the three-level learner-modeling framework for analyzing and representing learners' proficiency in cognitive skills and strategies for complex decision-making tasks (Rajendran, et al., 2017). The specific instance illustrated in the figure is derived for the UrbanSim tutor that we have implemented in the Generalized Intelligent Framework for Tutoring (GIFT) environment (Biswas, et al., 2017). The learner model makes a set of inferences based on the learners' observed behaviors relying on the learning environment to provide an evaluation of the users' performance on specified learning goals and tasks in the decision-making environment. In the lowest tier of the learner model, the list of observable actions that are linked to accomplishing a subtask or a task are represented as cognitive skills (Segedy, et al., 2015a). At the next level, cognitive strategies define how cognitive skills can be combined to accomplish more complex tasks and learning goals. At the next higher level, the learners' intent in using a skill or strategy, as well as their monitoring, evaluation, and reflection behaviors are represented as metacognitive processes. Each node in the learner model has associated performance measures that we have generically called *competence* and *trend*. Competence captures the users accumulated proficiency on a skill, strategy, or metacognitive process, and trend is a measure of how the proficiency changes over recent time intervals (this could be a time interval for a continuously evolving training scenario or a discrete set of turns in a turn-based game (Rajendran, et al., 2017)). Cognitive skills and strategies the learner employs are inferred from their interactions (learning behavior) with the environment (Tscholl, et al., 2016).

We have also proposed an instructional strategy framework to support adaptive scaffolding. Our approach reviews learners' proficiencies in the hierarchical learner model as they work on their complex decision-making and problem solving tasks to select appropriate skills and strategies for feedback to be provided to learners. We have used this framework to develop adaptive scaffolding on skills and strategies; however, inferring learners' metacognitive processes is a challenging task, since a lot of it happens in the learner's mind. The current learner model is implemented (Biswas, et al., 2017) for a counterinsurgency domain linked to the turn-based open-ended game environment called UrbanSim (McAlinden, et al., 2008). The multi-level hierarchical learner-modeling scheme was first proposed in Kinnebrew, et al. (2017), and it has been applied to K-12 STEM learning environments, such as Betty's Brain (Biswas, et al., 2016) and CTSM (Basu, et al., 2017).

DISCUSSION

In order to apply the multi-level learner model across multiple learning environments, user's proficiencies on the set of cognitive skills, cognitive strategies, and metacognitive processes that are pertinent to the domain of interest, for example, land navigation, will be represented as part of the generalized hierarchy. As a first step, to apply the learner-modeling standard, researchers need derive the task and skill structure that is relevant to the learning and problem-solving domain. For example, in system's like Betty's Brain, CTSiM, and UrbanSim we interpret the students actions in terms of three learning tasks: information acquisition, solution construction, and solution assessment (Segedy, Kinnebrew, & Biswas, 2015; Kinnebrew, Segedy, & Biswas, 2017; Rajendran, et al., 2017). For example, we provide the three categories of learning tasks for Map Reading and Land Navigation domain below.

- *Information Acquisition:* finding, reading, and interpreting map data; using instruments such as GPS, markers, and waypoints; interpreting sketches.

- *Solution Construction:* terrain association; dead reckoning; route planning; computing distance in different terrains; pacing; following handrails; checking attack points, and analyzing and keeping track of the resources needed.

- *Solution Assessment*: finding optimal routes; safety and risk analysis; evaluating plans and rerouting.

The hierarchical task model is created by mapping the specific skills required for each training scenario to a more general set of skills and concepts for learning tasks. This hierarchical structure will also help us establish the interoperability of skills across multiple problem solving scenarios, and across domains. Though the lower levels of the hierarchy may recognize contextualized skills for specific scenarios, the higher levels may capture abstractions of the skills that apply across different aspects of problem solving, and across different problem scenarios. This will serve two important purposes: (1) we can accumulate learners' skill competencies across multiple training environments to study their overall progress; and (2) as guidance and feedback, we can remind learners of skills they applied successfully in other scenarios, and provide hints on how the skill may be applied in the current scenario.

The cognitive strategies that we developed for the UrbanSim environment: situational awareness, situation assessment, tradeoff analysis and second and third order effects, are general and relevant to other complex decision-making tasks, such as land navigation, or providing emergency assistance. However, they may manifest as a combination of domain-specific skills and tasks that are particular to individual training scenarios. Metacognitive processes are by definition domain independent, but the challenge with metacognitive processes is to detect them – a lot of the metacognitive reasoning may happen in the trainees head, and it is hard to know how exactly they are evaluating their situations and monitoring their progress, without having a directed conversation with the trainee. To decide when to initiate a conversation with the learner, we will design the metacognitive model framework and develop an algorithm based on data-driven approaches (Kinnebrew, Segedy, & Biswas, 2017). As a generalization, we conjecture that as learning modeling incorporates higher order thinking and self-regulation processes, we may need to increase the modalities that support learner modeling, for example, speech (understanding), face tracking, and physiological measures that are linked to motivation and intent.

CONCLUSIONS AND RECOMMENDATIONS

It is clear that next generation learning environments need to be developed to support multiple learning and training applications, and be responsive to multiple input modalities that include text, speech, physiological sensors, a video input that provides more detailed information about the learner interactions and their affect (emotions), and eventually even EEG data. This has multiple implications in terms of standardizing learning environment architectures, the nature of their interfaces, and in particular, the learner modeling schemes. The learner-modeling scheme must be rich and go beyond performance assessments, to capture behaviors (skills and strategies), metacognition, affect, motivation and intent to build a complete understanding of the learner in specific problem solving scenarios. Moreover, the overall scheme must generalize across different scenarios to be a truly acceptable and generalizable standard.

In this paper, we have presented a three-level hierarchical structure that captures cognitive skills, strategies, and metacognitive processes across a variety of learning scenarios that includes complex decision making tasks that apply across different problem domains. For the future, we propose to apply our framework across multiple learning environments to demonstrate its generalizability and its evolution to a standard architecture for learner modeling.

REFERENCES

Basu, S., Biswas, G., & Kinnebrew, J. S. (2017). Learner modeling for adaptive scaffolding in a computational thinking-based science learning environment. *User Modeling and User-Adapted Interaction, 27(1), 5-53*.

Biswas, G., Rajendran, R., & Mohammed, N. (2017) Developing a Metacognitive Tutoring Framework within the GIFT Tutoring Framework. *Final Project Report submitted to U.S. Army Research Laboratory –Human Research & Engineering Directorate*

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty's Brain. *International Journal of Artificial Intelligence in Education, 26(1), 350- 364.*

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta psychologica, 81(3), 211- 241*.

Hannafin, M. J., Hall, C., Land, S. M., & Hill, J. R. (1994). Learning in open-ended environments: Assumptions, methods, and implications. *Educational Technology, 34, 48-55.*

Kinnebrew, J. S., Segedy, J. R., & Biswas, G. (2017). Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Transactions on Learning Technologies, 10(2), 140-153*

McAlinden, R., Durlach, P. J., Lane, H. C., Gordon, A. S., & Hart, J. (2008). UrbanSim: A game-based instructional package for conducting counterinsurgency operations. *Army Research Development and Engineering Command - Simulation & Training Technology Center, Orlando FL.*

Rajendran, R., Mohammed, N., Biswas, G., Goldberg, B. S., & Sottilare, R. A (2017). Multi-level User Modeling in GIFT to Support Complex Learning Tasks. *In Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym5).*

Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2015). Using coherence analysis to characterize self-regulated learning behaviors in open-ended learning environments. *Journal of Learning Analytics, 2(1), 13-48.*

Segedy, J.R., Kinnebrew, J.S., Goldberg, B.S., Sottilare, R.A., & Biswas, G. (2015a). Designing Representations and Support for Metacognition in the Generalized Intelligent Framework for Tutoring. *In Proceedings of the 17th International Conference on Human-Computer Interaction. Los Angeles, CA, USA. In Foundations of Augmented Cognition (pp. 663-674). Springer International Publishing.*

Tscholl, M., Biswas, G., Goldberg, B. S., & Sottilare, R. A (2016). Detecting Metacognitive Strategies through Performance Analyses in Open-Ended Learning Environment. *In Chen, W. et al. (Ed.), Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education.*

Wood, R., & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of personality and social psychology, 56(3), 407*.

ABOUT THE AUTHORS

Dr. Gautam Biswas *is a Cornelius Vanderbilt Professor of Engineering, and a Professor of Computer Science, Computer Engineering, and Engineering Management in the EECS Department and a Senior Research Scientist at the Institute for Software Integrated Systems (ISIS) at Vanderbilt University.*

Dr. Ramkumar Rajendran *is a Postdoctoral Researcher at the Institute of Software Integrated Systems at Vanderbilt University.*

DISCUSSION COMMENTS

- ➢ Authors propose a task based hierarchical model that identifies proficiencies and then applies it in the domain; learner skills measured in one learning environment can be used to measure skills in another learning environment; performance of other skills build on other skills in other tasks; ability to master task → mastery of domain knowledge
- ➢ Dr. Goldberg mentioned *teaching through analogy* (if an experience is recorded from a previous task that they performed at skill level, it can be used as analogy in the next task to assist in other skills built on that other skill) – could be beneficial to this approach
- ➢ Mr. Barr pointed out a possible challenge regarding the extent of our assumptions of how far we learn; Dr. Graesser mentioned empirical data on prerequisite history and transfer as a candidate is dismal; Dr. Durlach suggested including measures of retention.

Bounding the Possibilities for Adaptive Intelligent Systems Standards and Implementation

Dr. Lauren Reinerman-Jones, Ms. Elizabeth Lameier, and Dr. Daniel Barber

University of Central Florida

Topical Area: Learner Modeling

INTRODUCTION

Adaptive Tutoring Systems (ATS) are rapidly developing and adaptions to personalize and optimize learning are expanding. Adaptations within ATS show boundless possibilities for augmenting learning and improving retention. Current best practices for a Learner Record Store (LRS) is using manageable chunks of information relevant to learning and adaptation. However, profiles needed for adaptation cannot encompass everything stored in an LRS. Baselining also cannot be solely based on a learner's knowledge level, and trend lines need to be sensitive to continuous monitoring of the learner to determine intervention invocation. Authors of ATS and researchers are multidisciplinary. Therefore, multi-functionality (i.e. across domain and context) calls for standardization. Increasing interdisciplinary efforts advances methods and strategies shared, however, each has their own technical jargon, theories, and frameworks. Standardization includes development of operational definitions, taxonomies, measurement of the effectiveness of adaptation, and determination of valuable data for the learner or instructor. A key challenge for standardization in ATS is a lack of process to arrive at comparable outcomes. Research to date is mainly standalone and case studies (Vandewaetere, Desmet, & Clarebout, 2011).

Standards Use-Case

To achieve the goal of standardization with ATS, it may help to examine the process followed in creating existing standards, in particular those supporting large communities and complex systems. The Joint Architecture for Unmanned Systems (JAUS) is a Society of Automotive Engineers (SAE) standard and an initiative started in 1998 by the U.S. Department of Defense (DoD) to develop an open architecture for unmanned systems (US) with the goal of interoperability, (Wikipedia, 2018; Society of Automotive Engineers International, 2012). US's are inherently complex, requiring expertise from a broad spectrum of professions and disciplines just to accomplish what may seem like a basic level of autonomy.

The first step in the process for creating JAUS was to recognize issues impacting the larger community. In particular, controls and sensors were highly customized to a specific US, and could not transfer to others, impacting procurement, training, and research. Next, leaders within the community, driven by customers (e.g. U.S. DoD), formed groups and committees to work the problem, resulting in the JAUS Working Group (WG) that included members from academia, industry, and government. This working group reviewed the state- of-the-art in US technologies to identify common functionality, practices, needs, and gaps. Through this process, US capabilities were broken down into components to the point where going any further would not make sense for a general architecture. For example, a "Global Pose Sensor" provides position and orientation data for a US. This data could be generated from a GPS, digital compass, or many other devices, but decomposing further than "Global Pose" would result in having to specify a technology, rather than a necessary functionality. Once these core services were identified, the JAUS WG was able to document a services purpose, functionality, interface, and data for use within an open architecture supporting a variety of cross-domain applications, (JAUS Working Group, 2007). Individual performers used this JAUS Reference Architecture (RA) for implementation within their own systems and eventual interoperability testing and demonstration across individual implementations. Assessments of these standardized tests showed that different vendor's hardware and "black box" systems could work together, and identified issues within the RA for further refinement.

Once a critical mass was reached with the JAUS RA documents, the JAUS WG migrated the effort to the Society of Automotive Engineers, Aerospace Division, Avionics Systems Division following established processes for publication and maintenance of JAUS. The AS4, Unmanned Systems Technical Committee now maintains and advances a set of standards documents, based off the original JAUS RA. This committee follows formalized procedures including reviews, timelines, and voting for approval of new documents and changes to old. The advancement of US capabilities is still ongoing, and therefore the standard is continuously expanding to support these services as they mature in academia and industry following the same processes for creating the original service set.

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

The term standards implies a community of interest seeking to baseline or normalize requirements to meet a greater goal. In research, that often means to enable systematic, yet evolutionary conclusions for the improvement of mankind or some other organism. In training and education, the goal is faster and better knowledge or skill acquisition. For Adaptive Intelligent Tutoring, the goal to achieve is yet undefined, but is probably a blend of the goals for research and training and education. The first charge, then, for the learning sciences community interested in ATS is to determine the goal that standards shall serve and identify the key areas to develop standards. The Center for Adaptive Instructional Sciences (CAIS) is serving as an outlet to do this.

The next step is to recognize the interdisciplinary nature that the learning sciences approach is bringing to bear on ATS. This community of interest is composed of instructional designers, educators, computer scientists, computer engineers, human factors researchers, and modeling and simulation practitioners.

Therefore, it is essential to operationalize key terms and constructs used widely, but often with different meaning from one person to the next. This sounds simple, but requires effort to understand the different perspectives or schools of thought each discipline approaches ATS problems. It also requires concession to not nit-pick every nuance of the definition, but compromise on shared general understanding of a concept. The standards themselves will help flesh-out some of the nuances.

Once a shared lexicon is established, then a common operating platform should be decided. This might be an actual platform, but certainly should contain foundational content. Currently, most efforts in ATS are built stand-alone such that functionality or lessons learned are not readily available to transition to another domain, learning strategy, individual difference, adaptivity, and so on.

DISCUSSION

Developing a Goal for ATS Standards

It is simple to say that standards are needed, but setting an actionable goal for standards to achieve is more complicated. A starting point for developing a goal for ATS standards is to consider that standards shall encompass functionality and a means to develop systematic, decisive conclusions. Standards at this level are anticipated to enable the following questions to be answered:

1. What are all the current ways to assess, adapt, and collect data in an ATS?
2. What assessment, adaption, data, or combinations thereof should be used?
3. When should an ATS assess, adapt, or collect data?
4. Where does an ATS implement assessments, adaptations, or data?
5. Who are the individuals that require the assessments, adaptations, or data for learning?
6. How does an ATS implement the assessment, adaptation, or monitoring of the data?
7. What level of individualization does an adaptive tutor need to remain for it to be effective and efficient?
8. Which assessments, adaptations, or data are most effective and meaningful to be made a priority for the ATS?

Developing a Shared Lexicon

Listening to a conversation between an Instructional Systems Designer (ISD) and an Educator about any training and education topic, one would think they were hearing two different languages. For example, the topic of implementation of a learning module is often approached by an ISD by referencing a rubric, whereas an educator focuses on the material and instructor connecting to the learner. This same topic would be further complicated by adding a computer engineer or human factors specialist into the conversation. Therefore, it seems obvious that establishing a shared lexicon with associated operational definitions is essential.

A good place to start operationalizing constructs is the beginning. The following are proposed operational definitions for the community of interest to discuss and come to agreement. An ATS is a training or education technology designed to adjust to an individual learners needs throughout a learning module or course. Adaption is the type of adjustment a system implements to meet the learners needs. Adaptivity is the ability of a system to adapt. Levels of Adaptivity has often been used to refer to Macro or Micro-Adaptation, whereby Macro-Adaption infers after a specific number of questions or following a complete learning module and Micro-Adaptation is intend to be more instantaneous such as following every question. Those three terms are still relatively vague, so a suggestion is to use Point of Adaptivity to refer to the instantiation of an adaption at either a point in time, after a specific assessment point, after a simulated lesson event, following a module, or any other instantiation of an adaptation. Assessment was described as collecting relevant data that is used for decision making, (Cooley & Lohnes, 1976). Some examples of assessments are formative assessments, summative assessments, assessments to measure effectiveness of the tutor, knowledge tracking, model tracing, algorithms, surveys/questionnaires, real-time physiological assessments and the use of each with differ depending on the context, domain, type, and administration (Koedinger, Anderson, Hadley, & Mark, 1997; Kulik & Fletcher, 2016). Each of these needs to be defined and should be expressed as relevant or irrelevant assessments to use for skill versus knowledge acquisition learning objectives. Skill acquisition is learning to perform an action, behavior, or process and can be physical or mental. Knowledge acquisition is learning of facts, ideas, and concepts for application or insight. Many other concepts need to be operationalized, but those discussed here are essential.

Developing a Standardized Platform and Content

Content is limited and content is very expensive (time consuming) to develop. To address this, one recommendation is to target a single topic and exacerbate all possibilities. If one topic is built out by many, then content will be more readily available and research questions can more efficiently, effectively, and conclusively be answered. Stated differently, to compare assessments, adaptations, and outcomes, content needs to be scoped to start with one type of learning, one domain, one individual difference, one learner strategy, and one set of adaptations to understand the range of manipulations and outcomes. Thus, standards in application for how and when to adapt, and for designing intelligent architectures, can be achieved.

CONCLUSIONS AND RECOMMENDATIONS

Establishing standards for ATS at first appearance may seem unmanageable, however there are structured processes the community can follow to bound the problem. JAUS is just one example of a successful standard resulting from a cross-disciplinary domain enabling not just interoperability possibilities for developers, but clear definitions and capabilities for customers to leverage in procurement activities. This paper proposes steps, based on existing best practices, for the ATS community to adopt to help government, industry, and academia conduct transformative instructional science. The best way for the community to accomplish this is to leverage what already works in other sciences: common lexicons, shared data, well defined system functionality, and benchmark tasks.

Many specialties within computer science (e.g., computer vision, machine learning) have established shared data sets to benchmark advancements in the field. For example, algorithm A is able to classify faces in an image set better than algorithm B in terms of accuracy and time. These benchmarks enable researchers to build upon the work of others with clear differentiation and replication of findings as they move towards a common goal. The ATS community must move in the same direction to avoid further one-off experiments on unique tasks that do not translate to other domains. Following the steps described here, working groups should be staged to address issues presented, and identifying commonalities across ATS to document best practices and interoperability of system components.

REFERENCES

Bagheri, M. (2015). Intelligent and adaptive tutoring systems: How to integrate learners. *International Journal of Education, 7*(2), 1-16.

Cooley, W. W., & Lohnes, P. R. (1976). *Evaluation research in education.* New York, NY: Irvington Publishers.

JAUS Working Group. (2007, June 27). Reference Architecture Specification. *2*. Retrieved from http://www.min.uc.edu/robotics/research/Final%20RA%20V3-3%20Part%201.doc

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-43. Retrieved from http://telearn.archives-ouvertes.fr.hal00197383/

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42-78.

Society of Automotive Engineers International. (2012, 08 16). *JAUS History and Domain Model.* Retrieved 02 23, 2018, from SAE International: https://www.sae.org/standards/content/air5664a/

Sonwalker, N. (2013). The first adaptive MOOC: A case study on pedagogy framework and scalable cloud architecture - Part I. *1(P)*, 22-29. New Rochelle, NY: Mary Ann Liebert, Inc.

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*(1), 118-130.

Wikipedia. (2018, 08 23). *JAUS*. Retrieved 02 09, 2018, from https://en.wikipedia.org/wiki/JAUS

ABOUT THE AUTHORS

Dr. Lauren Reinerman-Jones *is the Director of Prodigy, which is one lab at the University of Central Florida's Institute for Simulation and Training, focusing on assessment for explaining, predicting, and improving human performance and systems.*

Ms. Elizabeth Lameier *is a Research Associate in assessment, training, and education for the University of Central Florida's Institute for Simulation and Training Prodigy Lab. Previously, she was an Intervention Specialist and Early Childhood Educator at an inner city public school where she regularly adapted, modified, and individualized instruction, assessment, and monitoring in order to increase academic, social, and behavioral outcomes.*

Dr. Daniel Barber *is an Assistant Research Professor at the University of Central Florida's Institute for Simulation and Training. Dr. Barber has extensive experience in the fields of robotics and simulation, developing virtual plat- forms and tools for synchronization, processing, and streaming of data from physiological sensors within experimental and training environments supporting real-time adaptations to user state. His current research focus is on human system interaction and training assessment including multimodal communication, user interaction devices, teaming, decision-making, and adaptive systems.*

DISCUSSION COMMENTS

- The study group needs to construct a standardization goal for AIS and key areas to develop standards. Then, develop a process for arriving at standardization. JAUS serves as a user-case study for following a process or adopting their process. For more information, contact Dr. Barber (dbarber@ist.ucf.edu). He was a part of the JAUS committee.
- Authors focused on the need to operationalize key terms and constructs that are widely used and then develop a common operating platform/content that can be standardized; goal was to adopt an autonomous system of standards
- Prof. Graesser suggested documenting the major pressure points of problems among measurement between different stakeholders.

SECTION 5: DISCUSSION ON OPPORTUNITIES FOR PEDAGOGICAL MODELING STANDARDS

Opportunities for Pedagogical Modeling Standards

Dr. Paula Durlach

U.S. Army Research Laboratory

DISCUSSION COMMENTS

- Proposed standardization of either taxonomy or a framework of adaptive interventions. Dr. Sottilare pointed out that taxonomy would be useful because it provides an opportunity to use machine learning to help the authoring process. He also cited cognitive atlas as an example of how taxonomy is effective. Everyone concluded with the (3) major advantages of taxonomy in terms of: marketing, research, and policy makers.
- Prof. Graesser brought up a possible implementation of humans who are self-regulated (which is almost impossible). This is why being automated is important. However, why are teachers needed was the lingering question of the group. Dr. Sottilare pointed out that teachers are essential for workload management. Dr. DeFalco also noted that teachers are not monolithic. They possess a range of expertise. Therefore, expertise and adaption are critical in order for it to be truly adaptive. Being able to swap out context in the taxonomy is also crucial.
- Dr. Sommer brought up the reliability and validity of tools - usefulness, content validity, transferability, trajectory, past research; Prof. Graesser brought up the topic of transfer studies and its recent emergence. It initially consisted of descriptive statistics. The procedure is asking questions, then take a course. However, how useful can this material be in another similar course? Ms. Cockroft stated STEM as an example of transferability of content.
- How to go forward? Dr. Sottilare suggested using different interfaces, but same framework. Also, include a wide range of digital tools.
- How will it be applied? Dr. Sottilare suggested the need to think about different levels and what the differences between levels are – especially considering complexity. Prof. Graesser suggested researching standards in education for guidance.

Standardizing a Comprehensive Learning Ontology Repository with Pedagogical Identifiers

Dr. Jeanine A. DeFalco[1], Dr. Keith Brawner[2], Dr. S. Amy Sommer[3]

Army Research Laboratory, West Point, NY & Orlando, FL
[2]Army Research Laboratory, Orlando, FL
[3] United States Military Academy, West Point, NY

Topical Area: Authoring Tools

INTRODUCTION

The constantly evolving domain of adaptive training warrants a reexamination of standards for adaptive tutoring systems, specifically in the area of metadata tagging as it relates to the methods to curate, apply, and evaluate content in the authoring process. Metadata tags provide information not only to authors of tutoring systems, but are key in driving learning object searches. Metadata tagging represents key micro-communications between systems and search engines, and its absence limits the ability to meaningfully consume and exchange information with learning management systems. In the effort to standardize tools and methods for authoring content for generalizable tutoring systems, then, this paper will review a representative sample of tutoring systems and their metadata tagging methodology, then propose a more comprehensive ontological standard model that would include a pedagogical repository of scenarios with pedagogical identifier codes (PIC) (Hadji, Choi, & Jemnim, 2012), building on Koutsomitropoulos and Solomou's (2017) design for implementing a learning object ontology repository linked with thesauri datasets and sources.

REVIEW OF REPRESENTATIVE TUTORING SYSTEMS

Before offering a recommendation on best methodological practices on metadata tagging, a review was conducted to establish current practices on Intelligent Tutoring Systems (ITS) tag intervention and how content and interventions are mapped to tutors. The representative sample included practices done with ElectronixTutor (University of Memphis); Crystal Island (North Carolina State University); d2p Tutors (Pennsylvania State University); and CognitiveTutor (Carnegie Mellon).

ElectronixTutor

In an email exchange (2018), Dr. Art Graesser confirmed that ElectronixTutor directly tags knowledge components for content and learning resources for interventions. The pedagogy aspect is indirectly tagged at a macro level and explicitly tagged at a micro level. In terms of mapping content and interventions to tutors, there are current parameters that makes these links that are considered in the recommendation system, but the ElectronixTutor team is currently in the process of using standards to align tagging knowledge components for content and learning resources for interventions more comprehensively.

Crystal Island

In an email exchange (2018), Dr. Jonathan Rowe explained how the Crystal Island team has tagged some of Crystal Island's in-game reading passages (about microbiology content) with metadata about reading difficulty (e.g., Lexile score, mean sentence length, mean log word frequency, word count). At present, the system does not use these tags for anything specifically, but the data is available for future analysis. Additionally, in-game quiz questions are tagged with scoring information so then students can get immediate feedback by the system depending upon correct/incorrect responses.

Crystal Island had had a number of versions over the years, and the inclusion of tagged pedagogical interventions (high-, moderate-, and low detail information) have been in some versions. Dr. Rowe cited an example, where a clue or piece of in-game feedback from a non-player character would be tagged as high-, moderate-, or low-detail depending on the quantity of utterances provided by a character. Subsequently, an AI-based pedagogical model would be incorporated to make decisions about "what types of clues or feedback to deliver in different contexts using these tags to distinguish between the different types of feedback (i.e., should the tutor provide high-detail, moderate-detail, or low-detail feedback in this situation)" (Rowe, 2018; Rowe & Lester, 2015). Overall, metadata tagging has not been a core focus of Crystal Island, and their solutions have been admittedly ad hoc.

D2p Tutors

In an email exchange with Dr. Frank Ritter of Pennsylvania State University, the authors of this paper were informed that the d2p tutors are being tagged with content. However, Dr. Ritter informed that the developers of Tuples are considering using a complex AI/KA tool to tag pages.

Arizona State University

In an email exchange (2018), Dr. Van Lehn explained the four tutoring systems under development: Dragoon, TopoMath, Seatr/OPE, and FACT. As a part of the tutoring system, each have their own learning management systems (LMS) so metatags are not needed. A version of Dragoon works with 3 LMS (Moodle/GIFT, PAL3, Assessments). Communications use specific URLs that bypass metatags. The lessons are a sequence of activities in the tutoring systems and are JSON files. The description of activities are inside the JSON files and have knowledge components, which are tags for student learning.

Stottler Henke Associates, Inc.

In an email exchange (2018), Dr. Randy Jensen of Stottler Henke Associates, Inc. described the Follow Me InGEAR tutor. It is an operational tutor and authoring tool developed as a custom application with the Follow Me tactical game. However, being a standalone tutor, it was not integrated with an LMS or any larger learning architecture where it might fit within a predefined framework of learning objectives. While InGEAR is not implemented with metatags, InGEAR includes a catalog of 22 assessment measures, which correspond to learning objectives and can be incorporated into exercise scenarios. These assessments monitor the student's tactical decisions, and produce real- time interventions and debriefings.

Aside from the shorthand titles for assessment measures that are linked to interventions (feedback text), there is no meta-tagging infrastructure. Additionally, InGEAR does not use any formal tagging beyond naming the assessment measures and linking their objects to interventions implemented in the code. Moreover, the InGEAR system includes an authoring tool where interventions are specified and instantiated in exercise scenarios. (In the system's terminology, interventions involve assessment measures, and the corresponding feedback mechanisms and text).

Some pedagogical decision-making is supported in the authoring tool, such as the choice of interventions and the parametric values associated with them in scenarios. For example, the instructor/author can specify how many instances of a poorly chosen movement technique exhibited by the student during an exercise may be considered good, average, or poor. Note that the InGEAR assessment measures are designed for modularity, that is, for easy reuse in new scenarios using the authoring tool. So although there is no meta-tagging scheme, InGEAR supports generalization and reuse in a manner that can easily be compatible with such a scheme.

PROPOSED STANDARD: COMPREHENSIVE LEARNING OBJECT ONOTOLOGY REPOSITORY WITH PEDAGOGICAL IDENTIFIERS

Given the importance of metadata tagging but its inconsistent use and range of implementable tools, the authors of this paper suggest adopting and improving on Koutsomitropoulos and Solomou's (2017) proof-of-concept design to establish and maintain a semantics-aware learning repository. In this design, learning object (LO) metadata can be assigned machine-understandable semantic annotations by way of ontologies. In this way, the learning object ontology repository or (LOOR) can achieve an interoperability with other repositories and discovery tools at the semantic level.

Koutsomitropoulos and Solomou (2017) designed a learning object ontology based on the IEEE LOM standard and we demonstrated how LOs can be integrated with other semantic standards, like Simple Knowledge Organization System or (SKOS)-based terminologies to assist in description and discoverability. This resulting system has already been deployed to support online courses at the Democritus University of Thrace (DUTH), offering more than 1500 courses and labs, currently for two thematic disciplines: mathematics and medicine,

Further, Koutsomitropoulos and Solomou demonstrate how by linking resources to standardized thematic taxonomies and additional ontologies and datasets, LOs are then opened up to the linked data world and the Web of Data. As a proof-of-concept, Koutsomitropoulos and Solomou built on a web-based ontology management framework, WebProtege, and designed a process and services that allows its reuse as a LOOR. In their design and implementation of their learning object ontology repository (LOOR), manageability and discovery of LOs based on WebProtege were integrated with learning management systems where course authors can discover, point to, and reuse additionally learning material for their courses; share a pool of semantically enhanced LOs available for indexing and supported by collective intelligence; demonstrate the semantifcation of educational resources by exposing their LO metadata through ontologies; and finally demonstrate how, through other discovery mechanisms and digital repositories they link LOs with SKOS thematic terminologies.

However, for the purposes of the field of adaptive tutoring, an improvement to this learning object ontology repository would include expanding the functionality beyond themes and content objects to facilitate pedagogical annotations.

Pedagogical annotation according to Bloom's Revised Taxonomy

Hadji, Choi, and Jemni's (2012) paper presents a conceptual model of a courseware generation system that uses a pedagogical scenario model to support pedagogical flexibility in the adaptive courseware generation system. Notably, they compose pedagogical scenarios using a pedagogical identifier code that is given to reference each pedagogical scenario stored in a repository, including pedagogical objectives that identify the objective to achieve through the pedagogical scenario being modeled. Further, they identify the following pedagogical models which while not comprehensive are established examples: presentation, problem solving, discussion, brainstorming, games, simulation, role playing, case study, project design method, question and answer method. These models also include an activity-sequencing that can be used to apply a particular pedagogic scenario.

Essentially, the Hadji, Choi and Jemni (2012) model is a sound one, except the authors rely on the prior iteration of Bloom's taxonomy in which the ordering of cognitive skills was organized according to nouns, as opposed to the revised version. Bloom's Revised Taxonomy (Church, 2008) employs verbs rather than nouns, and swaps the highest of the thinking skills "evaluation" with "creation," – a meaningful and substantive change. In addition to relying on the original Bloom's taxonomy to develop their pedagogical scenario model, the authors correlate the order of the lower to higher thinking skills with novice to expert identifiers:

- Knowledge (Novice)
- Comprehension (Novice)
- Application (Average)
- Analysis (Average)
- Synthesis (Expert)
- Evaluation (Expert)

If one were to implement instead the Revised Bloom's Taxonomy, then, the correlation of novice to expert higher order thinking changes to:

- Remember (Novice)
- Understand (Novice)
- Apply (Average)
- Analyze (Average)
- Evaluate (Expert)
- Create (Expert)

These are important identifiers to include in a revised LOOR, because at a very clear and transparent manner, one can attribute not merely content to author adaptive courses, but more specifically identify content that will be effective in the overall adaptive learning process.

DISCUSSION

Standardizing a learning object ontology repository that would include expanding the functionality beyond themes and content objects, and include facilitate pedagogical annotations could be of great value to adaptive training course authors. Adapting Koutsomitropoulos and Solomou (2017)'s design to standardize a LOOR that can achieve interoperability with other repositories and discovery tools at the semantic level while also providing functionality to include pedagogical identifiers can aid in methods to curate, apply, and evaluate content for the authoring process

Arguably, learning objects and content are only as effective as the way in which they are implemented to promote learning and skill mastery. As such, not all objects are effective across all pedagogical frameworks. Therefore, including a pedagogical annotation function would allow pedagogical flexibility in the adaption process of creating courses.

It is important here to note that pedagogical tagging should include not just the way in which learning objects are presented to the learner, but rather note how the content supports critical thinking and discriminate intelligence. Therefore, encouraging the metadata tagging that identifies how learning objects support higher order thinking is important when scaffolding learning experiences that move the learner from novice to expert thinking, as aligned with Bloom's Revised Taxonomy.

CONCLUSIONS AND RECOMMENDATIONS

In all, the value of standardizing a revised, comprehensive LOOR that would include pedagogical identifiers is anchored in the fact that content, in and of itself, does not promote learning or mastery; rather, it is the way in which it is designed and implemented, and the outcomes one seeks to achieve with that content (higher order thinking skills; discriminate intelligence) that yields the greatest value to achieving robust learning outcomes and expert mastery. In sum, then, the value of establishing a standard where a semantically aware learning object ontology repository includes not only objects and concepts, but pedagogical identifiers allows for a more robust and flexible tool for authoring adaptive courses, particularly across a range of domains whose learning objectives and desired outcomes are not uniform – especially as it relates to adapting for novice, to journeyman, and expert learners.

REFERENCES

Brawner, K. (2015, August). Authoring Tools for Adaptive Training–An Overview of System Types and Taxonomy for Classification. In *International Conference on Augmented Cognition* (pp. 562-569). Springer, Cham.

Churches, A. (2008). Bloom's taxonomy blooms digitally. *Tech & Learning*, 1, 1-6.

Goldberg, B. (2015, August). Augmenting Instructional Practice in GIFT Using the

Engine for Management of Adaptive Pedagogy (EMAP). In *International Conference on Augmented Cognition* (pp. 595-604). Springer, Cham.

Graesser, A. (2018). Email correspondence with author DeFalco, J.

Hadji, H. B., Choi, H. J., & Jemni, M. (2012). Pedagogically-Driven Courseware Content Generation for Intelligent Tutoring Systems. *International Journal of Fuzzy Logic and Intelligent Systems*, 12(1), 77-85.

Koutsomitropoulos, D. A., & Solomou, G. D. (2017). A learning object ontology repository to support annotation and discovery of educational resources using semantic thesauri. *IFLA Journal*, 0340035217737559.

Marjanovic, O. (2007). Using process-oriented, sequencing educational technologies: some important pedagogical issues. *Computers in Human Behavior*, 23(6), 2742-2759.

Ochoa, X., Carrillo, G., & Cechinel, C. (2014, September). Use of a Semantic Learning Repository to Facilitate the Creation of Modern e-Learning Systems. In *Proceedings of the XV International Conference on Human Computer Interaction* (p. 92). ACM.

Ososky, S., Sottilare, R., Brawner, K., Long, R., & Graesser, A. (2015). *Authoring Tools and Methods for Adaptive Training and Education in Support of the US Army Learning Model: Research Outline* (No. ARL-SR-0339). US Army Research Laboratory Aberdeen Proving Ground, United States.

PwC EU Services. (2015). Methodology and tools for data standards governance and management for EU Institutions. Retrieved on 25 January 2018 at https://joinup.ec.europa.eu/sites/default/files/custom-page/attachment/methodology_and_tools_for_metadata_governance_and_management_for_eu_institutions.pdf

Ritter, F. (2018). Email correspondence with author DeFalco, J.

Rowe, J. (2018). Email correspondence with author DeFalco, J.

Rowe, J., & Lester, J. (2015). Improving Student Problem Solving in Narrative- Centered Learning Environments: A Modular Reinforcement Learning Framework. *Proceedings of the Seventeenth International Conference on Artificial Intelligence in Education*, pp. 419-428, Madrid, Spain.

Sánchez-Alonso, S., & Sicilia, M. A. (2005). Normative specifications of learning objects and learning processes: towards higher levels of automation in standardized e-learning. *International journal of instructional technology and distance learning*, 2(3), 3-12.

Sicilia, M. Á., Lytras, M. D., Sánchez-Alonso, S., García-Barriocanal, E., & Zapata-Ros,

M. (2011). Modeling instructional-design theories with ontologies: Using methods to check, generate and search learning designs. *Computers in Human Behavior,* 27(4), 1389-1398.

Sinatra, A., Brawner, K., & Goodwin, G. (2015). Considerations for learner modeling in a domain-independent intelligent tutoring system framework. Retrieved on 1/4/18 at

https://gifttutoring.org/attachments/download/1458/SinatraBrawnerGoodwin_DHSS2015_CameraReady_071515_FINAL.pdf

Sottilare, R. A., Long, R. A., & Goldberg, B. S. (2017, April). Enhancing the Experience Application Program Interface (xAPI) to Improve Domain Competency Modeling for Adaptive Instruction. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale* (pp. 265-268). ACM.

ABOUT THE AUTHORS

Dr. Jeanine A. DeFalco *is an Adaptive Training Research Scientist and Post-Doctoral Research Fellow for the Army Research Laboratory in Orlando, FL, through the Oak Ridge Associated Universities. Her current research is on examining the relationship between analogical and creative reasoning, and its effect on expert problem solving in critical care medical education. She received her PhD in Psychology in 2017 with a concentration in Cognitive Studies and Intelligent Technologies from Columbia University. Jeanine is currently working out of the United States Military Academy in West Point, New York.*

Dr. Keith Brawner *is a researcher within the Adaptive Training Team within the US Army Research Laboratory's Human Research and Engineering Directorate (ARL-HRED) in Orlando, Florida. He has 12 years of experience within US Army and Navy acquisition, development, and research agencies He holds an MS and PhD in Computer Engineering with a focus on Intelligent Systems and Machine Learning (ML) from the University of Central Florida (UCF). His current research is in ML, active and semi-supervised learning, ITS architectures and cognitive architectures. He manages research in adaptive training, semi/fully automated user tools for adaptive training content, and architectural programs toward next-generation training.*

Dr. S. Amy Sommer *is an Assistant Professor at the U.S. Military Academy at West Point, and was previously appointed at HEC Paris. She completed a Post-Doctoral Fellowship at Harvard in Social and Organizational Psychology, and PhD at Ivey Business School. Amy's research examines how organizational crisis and temporal dynamics influence on teams and leadership, to enhance learning, performance, and creativity. Her work is published in a number of sources including Nature Scientific Data, The Journal of Experimental and Social Psychology, and Group and Organization Management. Her research has received awards and is funded by various sources including three federal governments.*

DISCUSSION COMMENTS

- Dr. DeFalco proposes to further the LOOR anthology from Koutsomitropoulos and Solomous (2017) study on proof of concept design to establish and maintain a semantics-aware learning repository. The goal is to lead for expert problem solving. Coined the term "discriminate intelligence"
- Mr. Barr brought up the lessons learned from ADL architecture projects. He stated that every recommendation had its own pedagogy. This then lead to the question on how DeFalco intends to approach this. She responded that you don't have to share your content. Easy, hard, difficult is not meaningful – focus on the difference between good and better; Blooms taxonomy makes across domain content applicable.
- Dr. Sottilare recommended another model (DataShop) and further reading into NELL & NEIL.
- In conclusion, Dr. Sottilare mentioned synthetic environments with data tagging. These standards mimic environment with cultural inclusion (attributes) under an ISO standard called the Synthetic Environment Data Representation Specification (SEDRIS); see ISO/IEC 18023, 18024, 18025, 18026, 18040, and 18041 for more detail.

Thoughts on a Design Framework for Adaptive Instructional Systems

Dr. Paula J. Durlach

U. S. Army Research Lab

Topical Area: AIS Design and Interaction

INTRODUCTION

The meaning of "adaptive" in the context of instructional systems (where "instructional" covers both education and training) has been used in many different ways. For example the U. S. Army Learning Model 2020 – 2040 (TRADOC, 2017) uses it in a least three different ways to describe a learning environment.

- Ability to change/update content easily.
- Ability to provide content tailored to learner's experience and knowledge.
- Ability of the learner to tailor their own learning experience.

My definition of adaptive instructional systems (AIS) fits the second definition above. It presents the learner with experiences that keep him or her in a sweet spot for learning: not too easy nor too hard, challenging, but not frustrating. This sweet spot is something referred to as the zone of proximal development (ZPD), after Vygotsky (1978). Based on learner inputs, it diagnoses learner weaknesses or misconceptions and makes instructional interventions aimed at fixing these. This requires some information about the student (the learner model), either acquired through explicit questions, or inferred from learner interactions with the system. It also requires designed-in decisions about how to intervene and the content and rules required to do so. These requirements make the development of AIS more resource-intensive than one-size-fits-all interactive multimedia instruction (IMI). It would be nice if we knew, of all the possible types of interventions, which were most effective, and for these interventions, the most parsimonious learner model. In that case, we could make recommendations to learning developers about what to create for different types of learning to get the most return on in- vestment (ROI).

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

Providing an answer to developers is challenging not only because there is no agreed upon terminology for different methods of adaptation, but also because AIS differ in multiple design aspects. This makes identifying the relative effectiveness of different methods of adaptation problematic. A meta-analysis of step-based intelligent tutoring systems (Van Lehn, 2011) suggested that crucial to effectiveness is step-based feedback and support, combined with a mastery learning technique. But beyond this, there has been little systematic research on the individual design aspects of AIS that produce the best learning outcomes. The Framework for Instructional Technology (FIT—Durlach & Spain, 2014) was created to provide both a scaffold and a terminology to support instructional developers considering AIS alternatives to one-size-fits-all learning environments, as well as to support learning scientists aiming to investigate AIS effectiveness systematically. FIT does this by proposing four types of adaptation and five levels of adaptation for each type. It is based on the conviction that corrective feedback, support, and mastery learning are essential for an AIS to be effective. Breaking mastery learning into two decisions—what to do at the end of an activity when the learner has not yet reached mastery, and what to do at the end of an activity when the learner has achieved mastery - creates the four types of adaptation.

These two decisions are referred to as Micro-sequencing (pre- mastery) and Macro-sequencing (post-mastery). FIT adopts the conception of inner loop and outer loop behaviors described by (VanLehn, 2006); but, extends this to more than two loops. Figure 1 illustrates the loops of Feedback, Support, Micro-sequencing, and Macro-sequencing. Support and Feedback occur in the inner loop, micro-sequencing at the next loop, and macro-sequencing in the outer loop. There could be more outer loops, for example if courses sequence is also considered.

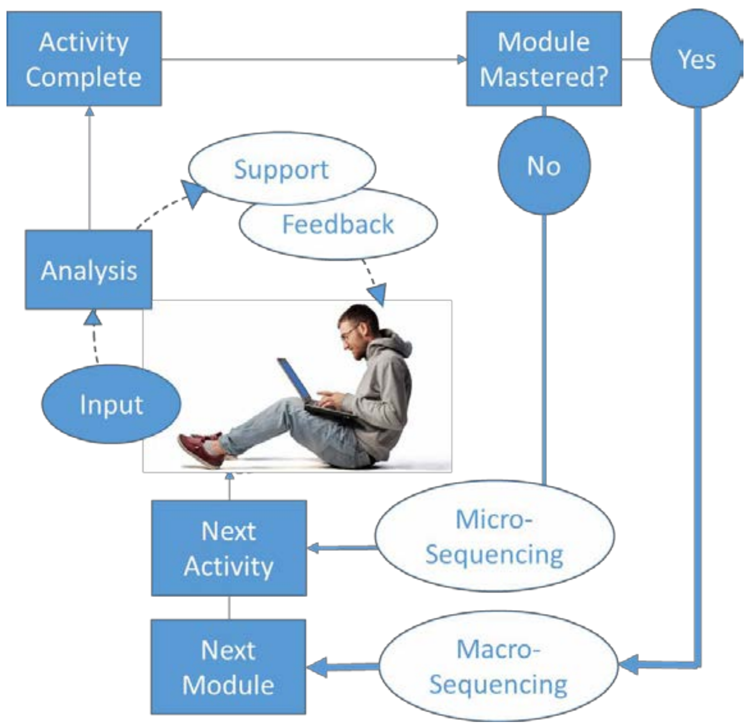


Figure 1. Illustration of where Support, Feedback, Micro-sequencing, and Macro-sequencing occur in AIS

FIT's levels of adaptation were intended to represent increasingly more personalized adaptations. For all of the four types of intervention, Level 0 represents little or no personalization. The different levels of adaptation for Corrective Feedback are listed in Table 1. At each successive level, error diagnosis and/or the information in the feedback message become successively more detailed. At Level 0, the student receives no personalized information, just a summary score on their performance. At Level I, the student is given item level information—what they did right and what they did wrong. At Level II the student is told not only what they did right or wrong, but when wrong, the correct response as well. At Level III (error sensitive) the feedback addresses the particular errors committed, drawing the student's attention to how their choice or action differs from the one designated as correct. Finally, Level IV (context-aware) corrective feedback additionally takes into account information about the student's past performance relevant to the knowledge components being assessed. Thus, different feedback might be given to a student known to have made the same error repeatedly vs. a student who has no history with the domain or concept. A complete explanation of each of the other three types of adaptations at each of the five different levels of adaptation is beyond the scope of this paper, but can be found in Durlach and Spain (2014), and a further explanation of how the levels are applied to Support can be found in Durlach (2014).

DISCUSSION

One of the motivations for attempting to define levels of adaptation was to provide an analogy with levels of IMI, with which training developers are already familiar (e.g., DoD MIL-HDBK 29612-3A, 2001; Strauss, et al. 2009). That schema lists four levels of interactivity as illustrated in Table 2.

MIL-HNDBK-29612-3A (2001) also presents three learning types, with each type also having levels. The types are Knowledge (5 levels), Skills (7 levels), and Attitudes (5 levels). The handbook then goes on to recommend which IMI level is appropriate for learning objectives at each learning type and level. The FIT model itself does not address learning type; but the FIT paper (Durlach & Spain, 2014) does make recommendations as to what level of IMI might be appropriate for different FIT adaptive level and adaptive type combinations. So one question to consider is whether a framework for AIS should take learning type into account.

Table 2. Levels of IMI

Level	Description
Level 1 - Passive.	The student acts solely as a receiver of information.
Level 2 - Limited participation.	The student makes simple responses to instructional cues.
Level 3 - Complex participation.	The student makes a variety of responses using varied techniques in response to instructional cues.
Level 4 - Real-time participation.	The student is directly involved in a life-like set of complex cues and responses.

One distinction implicit in FIT is that between local and context-based adaptation. Local adaptations use student input from one particular point in time or assessment event. They only take into account the last thing the student did. In contrast, context-based adaptations are defined as those that take into account other stored information about the student derived from previous interactions or other sources. E.g., Corrective Feedback Levels 0 – III as characterized by FIT are all local; only Level IV is contextual. FIT confounds the local/contextual distinction with level of personalization, however; it assumes that contextual is always more personalized than local; but this need not be. If a student receives the message “75% Great Job!” because they were previously failing, but an A- student gets, “75%, What’s Up?” is both Level 0 (summary score) and Level IV (contextual)? An improvement to FIT might be to take granularity of the intervention explicitly, as per see VanLehn (2011). Table 3 presents a potential revision of the original FIT Corrective Feedback characterization by crossing student data used and level of feedback granularity. Note that the boxes are not mutually exclusive. E.g., local problem level feedback might be combined with contextual summary feedback. Also note, there are multiple ways that a particular student data-by-granularity combination might be implemented as represented in the local problem-level and step-level boxes. Finally note, this framework does not represent the kind of data used (e.g., past performance, demographics, traits, states). So another issue for consideration is whether a taxonomy of adaptive interventions need take the nature of the learner data used into account.

Table 3. Potential alternative scheme of classification

		Granularity		
		Summary	**Problem Level**	**Step Level**
Data Used	**None**	No Feedback	No Feedback	No Feedback
	Local	Summary score (70%, Pass, Good)	Item accuracy Item accuracy + correct response Item accuracy	Item accuracy Item accuracy + correct response Item
	Contextual	Summary score and comment reference past performance or other comments based on other student data, such as traits	As per above with some reference to past performance or other comments based on other student data, such as traits or current affect	As per above with some reference to past performance or other comments based on other student data, such as traits or current affect

CONCLUSIONS AND RECOMMENDATIONS

As illustrated above, the more the framework attempts to represent all the decisions an AIS designer needs to make, the more complex the framework becomes. It may become so complex as to be difficult to use. At that point it is worth considering the purpose of the framework. When we constructed the FIT model we were trying to provide a guide to training developers on how to select the properties of AIS, by analogy to how the MIL-HNDBK tries to guide selection of IMI. But it turns out that selection of IMI by training developers (in the Army anyway), does not actually follow the handbook. Developers typically do not categorize individual learning objectives into one of the 17 learning type-by-learning levels specified in the handbook and assign the recommended IMI level to each learning objective. Rather, one of the four IMI levels is typically designated for an entire course, and this is often influenced by cost. Consequently we tried to limit FIT to two dimensions; but, as per the discussion, this limit leaves out important design considerations.

For AIS development teams and researchers a more complex framework might be acceptable. Were a framework sufficiently agreed upon, it might be integrated into a design tool to support AIS authors. Analogous tools have been created for making instructional design decisions (e.g., TARGET, http://bldr-webtest.alionscience.com/Target/), but the jury is still out on these kinds of tools (i.e., how much do they actually get used?). Perhaps the greatest usefulness of such a framework would be to guide research decisions on what types of adaptive interventions to compare to evaluate return on in- vestment. This paper proposes several dimensions for consideration such as: learning type (e.g. concepts vs. skills), learner data type (e.g., mastery level, affect), learner data recency (e.g., local vs. contextual), intervention granularity (e.g., problem-based vs. step-base), intervention type (e.g., feedback, support, remediation), and interactivity level.

REFERENCES

DoD MIL-HDBK-29612-3A (2001). Development of Interactive Multimedia Instruction (IMI), Part 3. http://www.navair.navy.mil/nawctsd/Resources/Library/Acqguide/296123a.pdf

Durlach, P. J. (2014). Support in a Framework for Instructional Technology. In R A. Sottilare, A. C. Graesser, X. Hu, & B. S. Goldberg (Eds.), *Design Recommendations for Intelligent Tutoring Systems, Volume 2, Instructional Management* (pp. 297 – 310). US Army Research Laboratory. https://www.gifttutoring.org/doc- uments/44.

Durlach, P. J. & Spain, R. D. (2014). Framework for Instructional Technology: Methods of Implementing Adaptive Training and Education. Technical Report 1335. U.S. Army Research Institute for the Behavioral and Social Sciences. www.dtic.mil/docs/citations/ADA597411.

Straus, S. G., Shanley, M. G., Burns, R. M., Waite, A., & Crowley, J. C. (2009). Improving the Army's Assessment of Interactive Multimedia Instruction Courseware. The RAND Corporation. https://www.rand.org/con-tent/dam/rand/pubs/monographs/2009/RAND_MG865.pdf.

TRADOC (2017). Army Pamphlet 525-8-2, U. S. Army Learning Model 2020 – 2040. http://www.tra-doc.army.mil/tpubs/pams/tp525-8-2.pdf.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, *16*, 227-265.

VanLehn, K. (2011): The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutor- ing Systems, Educational Psychologist, 46:4, 197-221.

Vygotsky, L.S. (1978). Mind in Society: The development of higher psychology processes. Cambridge MA: Har- vard University press.

ABOUT THE AUTHOR

Dr. Paula J. Durlach *has conducted and managed adaptive training research, as well as other projects, at U. S. Army and DoD organizations for more than ten years. She is currently at the U.S. Army Research Lab, Simulation and Training Technology Center, and can be contacted at Paula.J.Durlach.civ@mail.mil.*

SECTION 6: DISCUSSION ON OPPORTUNITIES FOR DOMAIN MODELING STANDARDS

Opportunities for Domain Modeling Standards

Dr. Xiangen Hu and Dr. Arthur Graesser

University of Memphis

DISCUSSION COMMENTS

- Authors proposed standardization of skills and knowledge sets, similar to Assessment and LEarning in Knowledge Spaces (ALEKS); once a learner knows a skill, and that learner has an existing skill then the combo of the two leads to learning of the third. Guarantees local model to global model. Knowledge space is gathered through SAT data, and then predicted probability to mastery tasks.
- Prof. Graesser pointed out how grain size seems daunting for this task and questioned the level of detail needed, and why. Xiangen replied that we don't know the level of detail but we do know the levels of domain but inside is difficult assumptions of homogeneous steps - need to revise specific learning beyond granularity; the content is not built to fit the model
- Dr. Fletcher questioned whether this proposed standard could be implemented with humanities. Xiangen explained that it could not due to the prerequisite logic.
- Dr. Graesser stated that the content standard doesn't tell how it is measured, or anything about the learner. A system would have to figure out the pedagogy, type of learner, etc.
- Mr. Cai clarified that the proposed standard is at a higher level than ALEKS, in particular how you specify domain knowledge. Mr. Barr brought up that prerequisites aren't addressed except what's hard or easy, what to do if they get something wrong.

Standardizing Content and Domain Models in Intelligent Tutoring Systems: A Few Ideas

Dr. Xiangen Hu [1,2], Dr. Arthur C. Graesser [1] and Mr. Zhiqiang Cai [1]

The University of Memphis [1], Central China Normal University [2]

Topical Area: Domain Modeling

INTRODUCTION

When considering standardization for each of the components of intelligent tutoring systems (ITS), domain models are perhaps the most visible and intuitive components to subject matter experts, curriculum developers, and learners. It is content and domain modeling that most subject matter experts think about when they create curricula and learning environments. The learners immediately explore the table of contents and ponder the relevance to their life and interest. The pedagogy, interface, and intelligent algorithms are occasionally considered, but are not as prominent as domain modeling. Stated most simply: "Content is king."

The challenge lies in the perception of many designers of learning environments that content and domain modeling runs the risk of being wide-open, arbitrarily constrained, potentially or typically ill-defined, impossible to organize in a well specified landscape, and often resistant to formal quantification. World knowledge is fraught with under-specification, vagueness, contradictions, and misconceptions. So what does it mean for establishing standards for content and domain modeling?

These challenges have not prevented researchers in cognitive science, semantics, computational linguistics, and computer science from attempting to precisify the mess and impose various forms of specificity, however fragile they may be. There are already widely used models for characterizing domains, such as "concept maps" (http://cmap.ihmc.us/docs/conceptmap.php), Knowledge Space Theory (http://liinwww.ira.uka.de/bibliography/Ai/knowledge.spaces.html), semantic taxonomic networks (such as the semantic web, Doan, Madhavan, Domingos, & Halevy, 2002; https://www.w3.org/2001/sw/wiki/Main_Page), the knowledge components of the Knowledge-Learning-Instruction framework (Koedinger, Corbett, & Perfetti, 2012; see also Graesser et al., in press), case-based scenario design (Jonassen & Hernandez-Serrano, 2002), and other proposals that were identified in the 4th volume of the *Design Recommendations of for Intelligent Tutoring Systems* that focused on Domain Modeling (Sottilare et al., 2016). These proposals need to accommodate differences between skills and knowledge associated with cognitive, psychomotor, teamwork, social, and other major categories of activities. Various taxonomies have been offered in different fields and are currently being explored in this arm of AIS; suggestions are welcome and have been offered by a number of AIS contributors (Bob Sottilare, Paula Durlach, Robbie Robson, and Avron Barr).

We believe it is important to identify these different frameworks or taxonomies and to see how they can be systematically coordinated with the other ITS components. For example, there have been attempts to relate knowledge components to pedagogical, interface, and computational algorithms (Koedinger et al., 2012; Durlach & Spain, 2014). It has been suggested that the common or frequent features of successful ITSs be identified and incorporated in the initial version of a standard for domain modeling. This suggestion of course presupposes that ITSs are sufficiently similar that common features would emerge.

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

When we consider standardization of a domain model for ITSs, there are a number of principles that we believe are essential:

- It should reflect the natural way that content developers think about the domain, as opposed to an obtuse model that is misaligned with their thinking.
- It needs to have a sufficient amount of mathematical, formal, or computational specification so that it is feasible to implement in an ITS.
- It should be consistent with existing ways of specifying domains that have been vetted by relevant communities, including those mentioned above.
- It should be coordinated with other ITS components with a standard communication protocol.

Domains may range from being formally well-defined with closed world semantics (Knorr, Alferes, & Hitzler, 2011) to being ill-defined open worlds that are quantitatively modeled with statistical, high-dimensional, semantic spaces (Landauer, McNamara, Dennis, & Kintsch, 2007). Differences on this continuum no doubt have repercussions on other ITS components that need to be clarified.

DISCUSSION

Our goal is to develop a standardized domain model for ITSs that is implemented digitally with computers. It is reasonable to assume, for example, that most if not all domain models have structural and relational specifications. That is, there are standard ways to specify types of items and types of relations between items. The specification would have several levels, aspects, and grain sizes. Furthermore, such as standardization for domain models would include the claims below.

1. *The structure property of a domain*. Different content domains may have structures that differ algebraically or formally. For example, the typical prerequisite structure for a domain like *mathematics* may not be applicable to other domains such as *history*.
2. *The elements (singletons or groups) in a given domain*. The smallest elements or primitives in a domain need to be defined, as well as clusters or collections.
3. *The types of elements (knowledge components, skills) and associated relations*.
4. *The relationships between the elements (or clusters) in a domain.* The relations between any two elements need to be specified. There are different types of relations that may or may be ordered (i.e., qualitative and quantitative assumptions).
5. *The mappings between different types of domain models for different ITS components.*

The properties of a simple domain model are relatively easy to specify mathematically, at least for well-defined domains such as mathematics, semantic taxonomies, and spatial organization. However, there are ill-defined domains that do not have constraints that lend themselves specify elegant structural properties (such as latent semantic analysis, neural networks, dynamical systems, and some classes of Bayes networks). Moreover, there are complications of even well-structured domains that require additional aspects to be identified and defined. For example, event A may cause B, and B cause C, but the transitive inference of A cause C would not be true. Or A is a prerequisite for B, and B for C, but the transitivity of A being a prerequisite for C would be incorrect. Or there may be a decision preference of A over B and B over C, but not A over C. Relations are often context-dependent, so a relation is true in one context or scenario but not another. It is conceivable that an alternative conceptualization is sometimes a useful option to this structural-relational standardization of domain knowledge.

CONCLUSIONS AND RECOMMENDATIONS

Establishing a well-received standard for domain modeling will likely be a long process for an active research, development, and implementation community in ITS. To start, we may first consider some *case studies* on popular ITS systems in specific domains. For example, there are existing, well-developed systems in mathematics, such as ALEKS, Cognitive Tutor, and ASSISTments (Heffernan, N. & Heffernan, C., 2014). Is there a domain standard for these systems? If there is, what is it? To conduct analyses on these case studies, it may be prudent to distinguish a domain structure standard from a domain content standard. Perhaps a domain content standard is impractical at this point of standardization, whereas a domain structure standard is feasible for handling some prototypical ITS. Perhaps an adequate communication protocol among all components of standardization will make it feasible to handle a broad array of domains with very different constraints. We recommend that these different avenues be explored at the next phase standardization efforts.

REFERENCES

Doan, A., Madhavan, J., Domingos, P., & Halevy, A. (2002). Learning to Map Between Ontologies on the Semantic Web. In *Proceedings of the 11th International Conference on World Wide Web* (pp. 662–673). New York, NY, USA: ACM.

Durlach, P.J., & Spain, R.D. (2014). Framework for Instructional Technology: Methods of Implementing Adaptive Training and Education. Technical Report 1335, United States Army Research Institute for the Behavioral and Social Sciences, Ft. Belvoir, VA.

Graesser, A.C., Hu, X., Nye, B., VanLehn K., Kumar, R., Heffernan, C., Heffernan, N., Woolf, B., Olney, A.M.,

Rus, V., Andrasik, F., Pavlik, P., Cai, Z., Wetzel, J., Morgan, B., Hampton, A.J., Lippert, A.M., Wang, L., Chen, Q., Vinson IV, J.E., Kelly, C.N., McGlown, C., Majmudar, C.A., Morshed, B., and Baer, W., (in press). ElectronixTutor: An intelligent tutoring system with multiple learning resources for electronics. *International Journal of STEM Education.* DOI 10.1186/s40594-017-0072-5.

Heffernan, N. & Heffernan, C. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. International Journal of Artificial Intelligence in Education. 24, 470-497

Jonassen, D. H., & Hernandez-Serrano, J. (2002). Case-based reasoning and instructional design: Using stories to support problem solving. *Educational Technology Research and Development*, *50*(2), 65–77.

Knorr, M., Alferes, J. J., & Hitzler, P. (2011). Local closed world reasoning with description logics under the well-founded semantics. *Artificial Intelligence*, *175*(9), 1528–1554.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, *36*(5), 757–798.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). (Eds.). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.

Sottilare, R. A., Graesser, A. C., Hu, X., Olney, A., Nye, B., & Sinatra, A. M. (2016). *Design Recommendations for Intelligent Tutoring Systems: Volume 4 - Domain Modeling*. US Army Research Laboratory. ISBN 978-0-9893923-9-6.

ABOUT THE AUTHORS

Dr. Xiangen Hu. *Professor of psychology, adjunct professor of computer sciences, electrical and computer engineering at the University of Memphis. Professor and dean, school of psychology, Central China Normal University.*

Dr. Art Graesser. *Professor of psychology at the Institute for Intelligent Systems at the University of Memphis. Honorary research scientist in the department of education at Oxford University.*

Mr. Zhiqiang Cai. *Research Assistant professor with the Institute for Intelligent Systems at the University of Memphis.*

The Granularity of Adaptive Instructions: Exploring Seven Orders of Time Bands through Learning Curves

Dr. Jong W. Kim

Army Research Laboratory

Topical Area: Domain and Learner Modeling

INTRODUCTION

This paper seeks to provide a pointer toward Instructional Modeling Standards. Particularly, it is of interest to consider the granularity of adaptivity level. For example, a cognitive modeling effort has provided useful information about the domain and the learner models (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995), and about assessment of performance changes—i.e., learning and retention (e.g., Anderson, Fincham, & Douglass, 1999). The most useful lessons learned would be the details of the skill development processes, so that an adaptive instructional strategy can be implemented in an adaptive instructional system (e.g., Koedinger, Anderson, Hadley, & Mark, 1997). However, years of the efforts are mainly restricted to cognitive tasks. Scholars, thus, wanted to embody such cognitive models (e.g., Ritter & Young, 2001) in an attempt to have a greater understanding of human behavior. The reason is that different domain models (i.e., cognitive, psychomotor, and social domain) would require different assessment measures and adaptive strategies in learning. Army Research Laboratory, thus, has initiated an investigation into the adaptive instruction psychomotor tasks beyond the desktop environment (Goldberg, Amburn, Ragusa, & Chen, 2017; Sottilare, Hackett, Pike, & LaViola, 2016). In this paper, based on the aforementioned efforts in cognitive and psychomotor domains, it is sought to provide a consensus understanding of granular adaptivity in instructions.

Performance curves are useful to visualize performance changes, and evaluate adaptive instructional systems (AISs). Particularly, in a formative assessment, where a student is being taught about a concept, a fact, or a task, an AIS should appropriately assess the student's learning progress, and properly inform the student of personalized and instructional contents. This is a very challenging and hard issue, but essential for advancing the learning experience.

As an assessment of performance changes, comparisons of two scores is a simple method. The two scores would be pre- and post-test results by comparing two learning systems, or by comparing before and after a unit of work, and by comparing scores after a period of time. This is called a summative assessment. In the meantime, the learner can be assessed *during a course of learning*, rather than at the end of a course, which helps us to see the skill development progress. The progress (i.e., an improvement by deliberate practice) could be visualized, and summarized in a *learning curve*. Learning curves have been used in industry as well in an attempt to investigate the prediction time (or cost) to produce a product (Jaber, 2016). In Cognitive Science and Education, learning curves have played a role to investigate practice effects by a certain knowledge representation in human memory systems (e.g., Anderson, Fincham, & Douglass, 1999). In both cases, learning curves follow a log-linear model (Newell & Rosenbloom, 1981).

AN OPPORTUNITY FOR STANDARDIZATION

Learning curves, however, seem not to be relevant to *making a standard* for the design of AISs since learning curves vary. Each learner would have different motivations and prior knowledge. Knowledge and skills are also varied and different. The task difficulty is diverse, and so forth. Indeed, it is diverse, but a standard way to look at the learning curve is necessary in an attempt to provide an understanding toward "Instructional Modeling Standard". The reason is that it can tell us about the grain size of adaptivity levels, and it would be useful for the AIS community to better provide ITS design alternatives.

Learning curves can provide important insights of how to define adaptivity levels. Educational outcomes generally reside in a certain time band of hours, months, and years. In the meantime, outcomes in a cognitive learning theory would be in the time band of seconds or milliseconds. Thus, regarding a meaningful assessment in training, should we look at the time band ranging from milliseconds to years? It is mentioned that there is a significant gap in an analysis of performance changes (Anderson, 2002), and the gap can be defined by Newell's time scale of human action (Newell, 1990, p. 122). We need the appropriate level of granularity and relevant theories. If we have a standard of how to look at performance changes with the grain size in the time band, it would be useful for the Intelligent Tutoring System (ITS) community to better provide instructional design choices.

In general, the task completion time follows a power law of learning representing a speed-up effect. Assessment of performance can be ranging from milliseconds to years; spanning of seven orders of magnitude (Anderson, 2002). This is useful in designing an adaptive instructional system since a large task can be meaningfully decomposed. One of the known pitfalls of learning curves is that a larger domain model or a large student sample size is likely to exhibit a better fit than a smaller one, even if the system does not teach the students any better (Martin, Mitrovic, Koedinger, & Mathan, 2011). For example, a larger task can be decomposed, but subtasks would be learned differently (Kim & Ritter, 2016). Thus, a simple comparison of learning curve in a large task seems not sufficient enough to make an instructional strategy. Furthermore, a near-term assessment by comparing learning curves would not be related to the long-term stability of learning (personal communication, Pavlik Jr, July 31, 2017).

Then, is there a reliable solution (as a standard) of how to compare learning curves in an attempt to provide improved adaptive instructions? This short paper brings up important considerations of the learning curve as an assessment tool in AISs. Particularly, we note an approach of using a Bayesian hierarchical model to compare learning curves by individuals and tasks/subtasks represented in a domain and student model. This would contribute to the standardized design of an ITS with a better understanding of short or long term stability of knowledge and skill learning.

DEFINING ADAPTIVITY LEVEL

Decomposing the Task

In Psychology, there is a decomposable hypothesis (Lee & Anderson, 2001). Simply, it indicates a larger task can be meaningfully decomposed to smaller unit tasks. Thus, a subtask (in seconds) might be relevant to educational outcomes. An experimentation that investigates human attention in milliseconds provides meaningful implications toward a longer time duration of a task. The adaptivity level would be related to how to decompose a task into smaller tasks. Therefore, a task can be meaningfully decomposed to a smaller unit of subtasks. A large, complex military task can be decomposed to cognitive subtasks, psychomotor subtasks, and physiological control subtasks (e.g., a tactical breathing subtask).

Time Bands

To define the adaptivity levels in an AIS, we need to add some kind of a time scale. I would like to adopt the best practice from other domain. It is Allen Newell's seminal work on Unified Theories of Cognition, and known as Time Scale of Human Action (Anderson, 2002; Newell, 1990). Newell first mentioned the

Band: Biological, Cognitive, Rational, and Social Band. 100 milliseconds to 100 hrs, even we think 10000 hrs of deliberate practice to be an expert (i.e., a surgeon performing a minimally invasive training). It inspired the efforts in the AIS community, and is seen in John Anderson's and Ken Koedinger's work.

It has been pointed out that there is a gap between millisecond level experimentations in cognitive psychology and months/years outcomes in education (Anderson, 2002). Anderson made a suggestion to use Newell's time scale of human action. In terms of his arguments, the learner model would reside in Newell's Cognitive Band (100 ms to 10 s), and the educational outcome lies in Social Band ranging from days to months. He argues using a cognitive modeling approach can bridge the gap in terms of a Decomposable Hypothesis, which suggests a large task can be decomposed to a smaller unit task.

Koedinger and his colleagues proposed a cognitive science based framework, called knowledge-learning-instruction (KLI), in an attempt to promote high potential principles for generality (Koedinger, Corbett, & Perfetti, 2012). He mentioned also Newell's time band of human action. Firing one production rule is assumed to take 500 ms (.5 s). Cognitive Tutor uses a set of production rules as a learner model. Reading time can be modeled using a set of production rules, and can be plotted in a log-log scale, representing a power law of learning. Educational outcomes would lie in Social Band with days to months. The granularity of adaptive instruction can go down to Cognitive Band with a production rule in ms. It is claimed that cognitive modeling can provide a basis for bridging between events on the small scale and desired outcomes on the larger scale.

The understanding of the time band (as a standard way to look at the granularity) in AISs can be instantiated in the domain and the learner model. Furthermore, the way to author the domain and learner model would be improved to better address the grain size of adaptivity.

AUTHORING AND ASSESSING: THE DOMAIN AND LEANER

Assessment of the learner's performance and corresponding measures would be different in terms of the domain. As a step forward, the understanding of the time band would be useful to deal with how adaptively the system behaves by specifying the different characteristics of domains. Different domains would include a cognitive domain, a psychomotor domain, and a social domain. Is, then, authoring one of such domains different from one another? How can we author a domain model in different domains (cognitive, psycho-motor, and social domains)? Are the learning curves different by different domain models? How are they different? Do they affect adaptive instruction as well?

Authoring a domain and learner model would be manual with or without a support from data mining and machine learning algorithms. The domain and learner model can be comprised of knowledge components. A manual authoring knowledge components approach to learner modeling. One kind of the knowledge component model is rule-based cognitive models that can be run through a computer simulation. Production rule-based models help in thinking about what knowledge may be needed to perform a particular task, how that knowledge might be decomposed to capture what the learner would do, and how widely specific knowledge components will transfer (Aleven & Koedinger, 2013).

The knowledge components based learner (domain) models can shape the behavior of an adaptive instructional system. Maintaining and updating the knowledge components based learner (domain) model would play a considerable role in the design of the grain size of adaptivity. Authoring tools can provide mechanisms to maintain and update the knowledge components based learner models, and to tailor aspects of the instruction (Aleven & Koedinger, 2013). Examples of intelligent tutoring systems that have taken a knowledge component (KC) approach to learner modeling are Cognitive Tutors (Anderson, Corbett, Koedinger, & Pelletier, 1995), constraint-based tutors (Mitrovic, Mayo, Suraweera, & Martin, 2001), and Andes (Vanlehn et al., 2005). Now, another research thrust is how to implement an automated approach to creating and refining such knowledge component based models. Variations in such models would be useful to understand the adaptivity.

DISCUSSION

Recently, I have been investigating skill development in a psychomotor domain, a golf putting, which would be a task that takes several seconds to complete. But, it would require days, months, or years of practice to see a satisfactory outcome. To help the skill development, it would be, therefore, necessary to consider the grain size of adaptive instructions and feedback. As mentioned earlier, a golf putting task can be decomposed into a cognitive subtask, a motor subtask, a subtask about controlling breathing (physiologically related subtask). Learning curves from all these subtasks would vary. In an initial ACT-R model of golf putting, the number of production rules is around 20. Learning in this domain involves the acquisition of such production rules. The production rule type knowledge components is useful, and this model can trace to find some sequence of productions that produces the behavior exhibited by the learner. A physio- cognitive model has been, recently, proposed to represent the slow breathing in a psychomotor task (Dancy & Kim, 2018). The physio-cognitive model can provide a rich understanding of the task for adaptive instructions. Again, it is necessary to propose a timing standard to look at performance—physiological control in seconds, cognitive thought process in seconds, and deliberate practice and outcomes in days.

CONCLUSIONS AND RECOMMENDATIONS

A time unit is used across the world as a standard, which helps us to communicate with each other regarding the time. A unit for the weight measure can be also another example. At least, we can successfully translate one unit to another and communicate with each other regardless of different views. Similarly, we could propose a standard regarding the granularity of adaptive instructions in an intelligent tutoring system. This would help us to better understand performance changes and their assessments. This would also help us achieve acceleration of learning by maximizing learning gains in a standard and objective way.

Newell's (1990) time band could be one candidate. There are several attempts that verify its usefulness for an intelligent tutoring system (Anderson, 2002; Koedinger, Corbett, & Perfetti, 2012). Based on this, an adaptive instructional system would be cognitively and physiologically inspired to produce instructional adaptivity. Adaptations in the time band could provide an understanding of a finer granularity of adaptive instructions. This effort would be helpful to achieve a cognitive training and brain plasticity (e.g., Fu, Lee, Boot, & Kramer, 2013) in an ITS.

REFERENCES

Aleven, V., & Koedinger, K. R. (2013). Knowledge component (KC) approaches to learner modeling. In *Design Recommendations for Intelligent Tutoring Systems - Learner Modeling* (Vol. 1, pp. 165-182). Orlando, FL: U.S. Army Research Laboratory.

Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science, 26*, 85-112.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4*(2), 167-207.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1999). Practice and retention: A unifying analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(5), 1120-1136.

Dancy, C., & Kim, J. W. (2018). Towards a physio-cognitive model of slow-breathing.

Fu, W.-T., Lee, H., Boot, W. R., & Kramer, A. F. (2013). Bridging across cognitive training and brain plasticity: A neurally-inspired computational model of interactive skill learning. *WIREs Cognitive Science, 4*(2), 225- 236.

Goldberg, B., Amburn, C., Ragusa, C., & Chen, D.-W. (2017). Modeling expert behavior in support of an adaptive psychomotor training environment: A marksmanship use case. *International Journal of Artificial Intelligence in Education*, 1-31.

Jaber, M. Y. (2016). Learning curves: Theory, models, and applications: CRC Press.

Kim, J. W., & Ritter, F. E. (2016). Microgenetic analysis of learning a task: Its implications to cognitive modeling. In F. E. Ritter & D. Reitter (Eds.), *Proceedings of the 14th International Conference on Cognitive Modeling* (pp. 21-26). University Park, PA: Penn State.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-43.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36*(5), 757-798.

Lee, F. J., & Anderson, J. R. (2001). Does learning a complex task have to be complex? A study in learning decomposition. *Cognitive Psychology, 42*(3), 267-316.

Martin, B., Mitrovic, A., Koedinger, K. R., & Mathan, S. (2011). Evaluating and improving adaptive educational systems with learning curves. *User Modeling and User-Adapted Interaction*, *21*(3), 249-283.

Mitrovic, A., Mayo, M., Suraweera, P., & Martin, B. (2001). Constraint-based tutors: a success story. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems* (pp. 931-940). Springer.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Lawrence Erlbaum.

Pavlik Jr, P. I. (July 31, 2017). Assessment of Forgetting: Enabling an ITS to Assess Subtasks Forgetting for an Adaptive Instructional Strategy.

Ritter, F. E., & Young, R. M. (2001). Embodied models as simulated users: Introduction to this special issue on using cognitive models to improve interface design. *International Journal of Human-Computer Studies, 55*(1), 1-14.

Sottilare, R., Hackett, M., Pike, W., & LaViola, J. (2016). Adaptive instruction for medical training in the psychomotor domain. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 1-13.

Vanlehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education*, *15*(3), 147-204.

ABOUT THE AUTHOR

Dr. Jong W. Kim *is an ORAU post-doctoral researcher, and an adaptive tutoring scientist at US Army Research Laboratory.*

DISCUSSION COMMENTS

- Author proposes standardizing learning curves to efficiently organize knowledge and skills within a complex task; mentioned decomposable hypothesis and Bayesian type models; need to optimize learning curve to provide adaptation for the learner; go from Stage 1 to Stage 3
- Discussion focused on suggestions and things to consider including letting the learner see the learning curve and possible skill decay to help with self-regulated learning, might want to consider other things that will affect learning and ability to perform tasks such as sleep or hunger, also consider what the model cannot capture such as regression periods or interference materials; informal learning affects also mentioned

The Potential of the National Information Exchange Model (NIEM) for Standardizing Adaptive Instructional System Information Exchange and Domain Knowledge

Dr. Paula J. Durlach

U. S. Army Research Lab

Topical Area: Interoperability

INTRODUCTION

Service-based, interoperable, open architectures is a significant gap in current distributed learning capabilities (The Advanced Distributed Learning, 2017). Addressing this gap would support the ability to share data by exposing it using modern web services or application programming interfaces (APIs). Such data could include learner performance data, usage data, learning content, and learning content metadata. Un- locking current data silos and sharing data across learning applications and systems could improve training, education, and human talent management. It could provide the basis for career or life-long personal learning assistants, and self-paced competency-based educational credentialing. It could also enable analysis of learning data which could be used to machine learn adaptive policies. Intelligent tutoring systems (ITS), while the epitome of adaptive instructional systems (AISs) are subject to this gap. They are designed to instruct on a fixed topic or skill set, and cannot go outside that, or reach out to the cloud to find prerequisite or more advanced content, or other ways of teaching the same content. Two ITS on the same subject (e.g., algebra), are unlikely to be able to exchange data, as they will typically have different domain and learner models (Brusilovsky, 2016; Lenat, 2016). This could be overcome if there were some standardization across systems. Robson and Barr (2013) discuss the benefits of ITS being able to "play well with others" via standardized methods of exchanging learner data. They also discuss the difficulties in establishing such standards, and suggest that efforts for standardization should focus on easing the integration of ITS into existing systems and sharing learner performance results.

PROBLEM OR OPPORTUNITY FOR STANDARDIZATION

The purpose of the present paper is to discuss an information exchange model, which may prove a useful basis for developing information sharing across not only adaptive instructional systems (AIS), but also across functional areas of an organization—e.g. to share learner records with human resources functions, or to use the knowledge management ontologies of an organization as the basis of a domain model for an AIS. Creating the domain model for an AIS is very resource intensive.

The idea proposed is that as more and more functional areas of an organization develop their own knowledge management models, these models may be repurposed as the starting points for creating AIS domain models for different functional areas. For example, if an Army organization already has developed ontologies describing tanks and different kinds of tank maneuvers, can those ontologies be used as part of a domain model for an AIS intended to teach tank maneuver? Additionally, if knowledge labeling conventions or ontologies are standardized within an organization, those conventions might be used as metadata for learning content, enabling discovery and recommendation linked to individual or team training needs. For example, can the tank and maneuver ontologies be adopted for labeling the learning objectives associated with content for teaching tank maneuvers?

In particular, this paper describes the National Information Exchange Model (NIEM, https://www.niem.gov/ as potential basis for standardizing AISs information exchange about learning objectives and learner data. NIEM is a reference model intended to enable information exchange among diverse public and private organizations. It can be thought of as a dictionary of agreed-upon terms, definitions, relationships, and formats, independent of how information is stored in individual systems. As illustrated in Figure 1, NIEM consists of core and domain-specific elements. The core includes common elements that domain-specific models build upon. For example, elements in the NIEM core include general concepts such as "person," "location," "item," "organization," and "activity." "Person" has about 200 subcomponents such as birth date, and sex. Domain communities of interest (CoI) build upon the core and can reuse elements already created by other communities to establish a model for their own needs. The concept of in- heritance is used throughout the NIEM data model to support extensibility. For example in the Emergency Management domain, "alarm" is a child of "activity" and therefore has all of the characteristics associated with "activity," as well as specific characteristics for "alarm," such as an alarm category code.

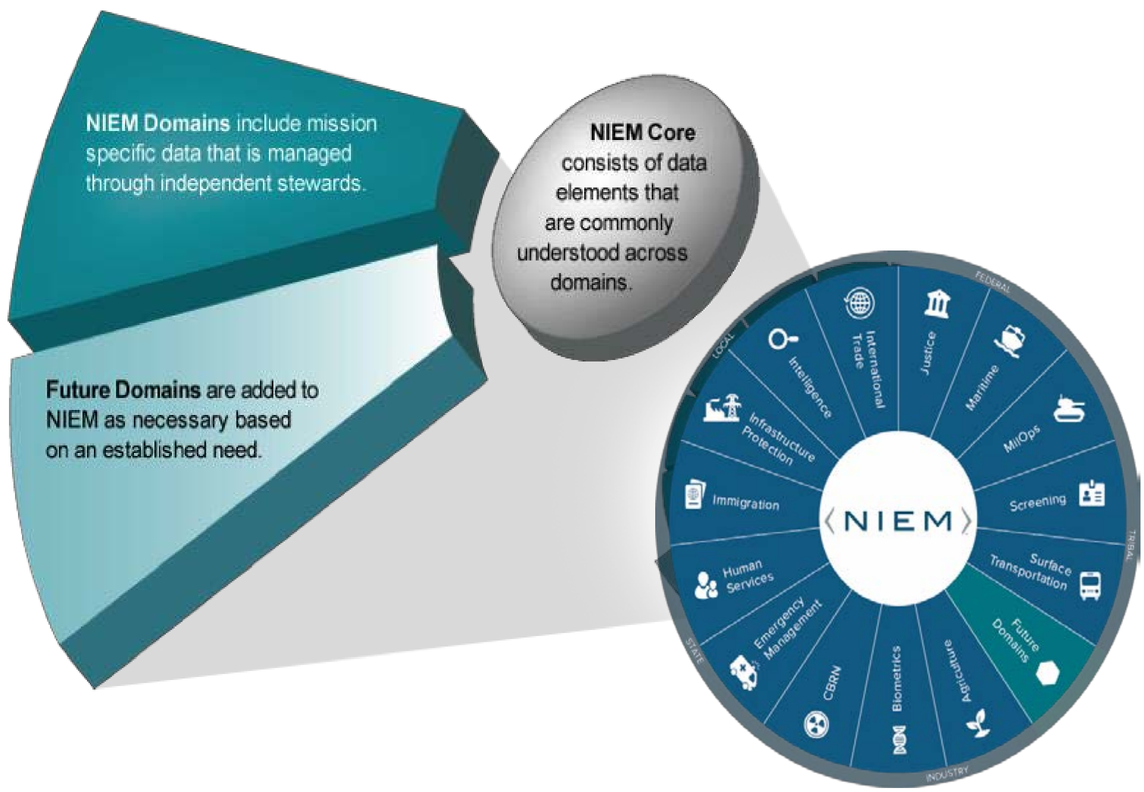


Figure 1. Illustration of NIEM Core and Domains

Domain communities of interest are formally established, to officially manage and govern a portion of the NIEM Data Model. An example of a NIEM domain is the Children, Youth, and Family Services (CYFS) domain. This domain supports timely, complete, accurate, and efficient information sharing to improve outcomes for children and youth whose circumstances make them particularly vulnerable. Within the CYFS domain, a governance committee stewards the domain and brings together CoI to identify information exchange business requirements. The CYFS CoI include but are not limited to: Juvenile Justice, Child Welfare, Child Support Enforcement, and Courts. End users across the communities develop and implement NIEM-based exchanges and provide new or updated information exchange requirements to the domain governance committee.

Information Exchange

Using NIEM, organizations are able to come together to agree not only on a common vocabulary, but also to use that vocabulary in a standardized data exchange via a NIEM information exchange package (IEP). The IEP is usually coupled with additional documentation, sample XML instances, business rules, and more to compose the Information Exchange Package Documentation (IEPD). An IEPD is the final product of the NIEM information exchange development process, also known as the IEPD Lifecycle. Using a core set of artifacts, the lifecycle consists of six phases: (1) planning, (2) requirements analysis, (3) exchange content model creation and mapping to the NIEM data model, (4) building and validating NIEM- conformant XML schemas, (5) package preparation into a single, self-contained, self-documented, portable archive file, and (6) publish and implement the IEPD for search, discover, and reuse. A NIEM-Unified Modeling Language (UML) is used to generate the NIEM-conformant exchanges. NIEM-UML can also help users manage NIEM domain data model content and create both a UML and XML schema representation. As illustrated in Figure 2, IEPDs can be reused—partially or fully—for a different needs within or across organizations.

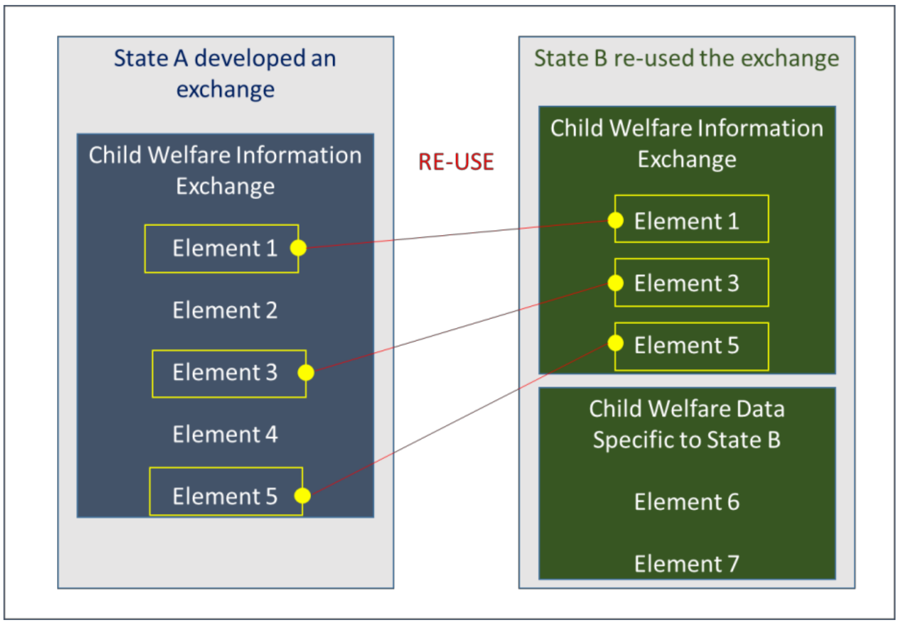


Figure 2. Illustration of IEP reuse.

National Electronic Interstate Compact Enterprise (NEICE) Success Story

The process to place children in homes across state lines was improved by using NIEM as the foundational data layer in an electronic system for information exchanges between states. This allowed disparate child welfare data systems to exchange information electronically. Within two to three weeks of launching the NEICE pilot, approvals that often took six to 12 months were taking one to two days, getting children into stable homes sooner. Savings from eliminating printing and mailing costs alone were close to $1.6 million. States can connect to NEICE in two ways, either through an automated interface that exchanges data between their IT system and NEICE Case Management System (which workers access via the web), or by directly connecting their IT system to NEICE Clearinghouse. Either way, states must translate their data into standardized fields by using this NEICE Information Exchange Package Document, which specifies the manner in which the automated data packets needs to be created. This automated data packet will be received by the receiving state child welfare IT system and ingested into their system for processing by case workers at the receiving state. Once the placement request is processed, the decision information is then sent back to the sending state for finalization of the placement. The NEICE Clearinghouse is a secure, cloud-based electronic exchange platform that allows one state's child welfare information system to send and receive data and documents to another state's child welfare information system. The Clearinghouse model allows states to process cases from within their state child welfare information system and avoid having to log into a different system to process cases. Participating states pay an annual service fee ($25,000) to access the NEICE system. This fee covers maintenance, server space, and administrative costs to keep NEICE operational 24/7, as well as ongoing enhancements to the system. As of December 2017, 24 states were using NEICE, 15 have plans to join, and eight are considering joining. More information and the IEPD can be found at https://www.acf.hhs.gov/completed-information-exchange-packet-documentation-iepd.

Security and Privacy

NIEM allows the use of metadata to describe specific requirements in regard to information security and the handling of sensitive privacy-protected information. Including this metadata allows systems that implement NIEM to automatically enforce rules that govern the use, protection, dissemination, and access controls for data being shared; however, other technologies are required upon exchange implementation to enforce security and privacy rules. This has been put to use in the Intelligence community, which established the Intelligence Community Information Security Marking (IC-ISM) as a standard for classified information.

DISCUSSION

Why NIEM instead of some other schema? NIEM is governed by a partnership of the U. S. Department of Justice, Department of Homeland Security, and Department of Health and Human Services. In 2013 the Chief Information Officer for the Department of Defense (DoD) issued the "NIEM first" memorandum, which directed that the DoD shall first consider NIEM for their data exchange standards. This was formalized in the 2014 DoD Instruction 8118.01, titled "Mission Partner Environment (MPE) Information Sharing Capability Implementation for the DoD and reiterated in the Army Data Strategy (Office of the Army Chief Information Officer/G-6, 2016), which has the stated objectives of making data visible, accessible, understandable, trusted, and interoperable. NIEM is currently being integrated into the NATO Core Data Framework (NCDF), and plans are afoot to integrate it into NATO a Standardization Agreements (STANAGs). Thus, it seems that NIEM is becoming widely adopted, and could be particularly useful, especially for integrating Army learner data with other functional areas of the organization.

As the Army, other services, and other government agencies move toward an integrated, service-oriented, information-sharing environment, the ability for data exchange should allow many users and applications to leverage the same data. Future technology-based training environments, such as the Army's planned Synthetic Training Environment (STE) are expected to operate within this context, relying on authoritative data sources for terrain and simulation models, including artificially intelligent computer-generated characters. AIS used to support training in the context of the STE would benefit from being able to exchange data with these functions, at a semantic level rather than having to interpret lower level simulation data and possess its own independent domain models. STE is expected to have the capability to automatically generate training scenarios appropriate for a unit's learning objectives. If these activities and objects were represented in NIEM (or some other machine readable and interoperable framework), the process of assembling a simulation-based training exercise, assessing the exercise, and recording the results could all be automated or semi-automated, as illustrated in Figure 3. The AIS policies and engine could use information about training needs and current skill levels to adapt the scenario appropriately (e.g. single threat vs. multiple threats or day vs. night.

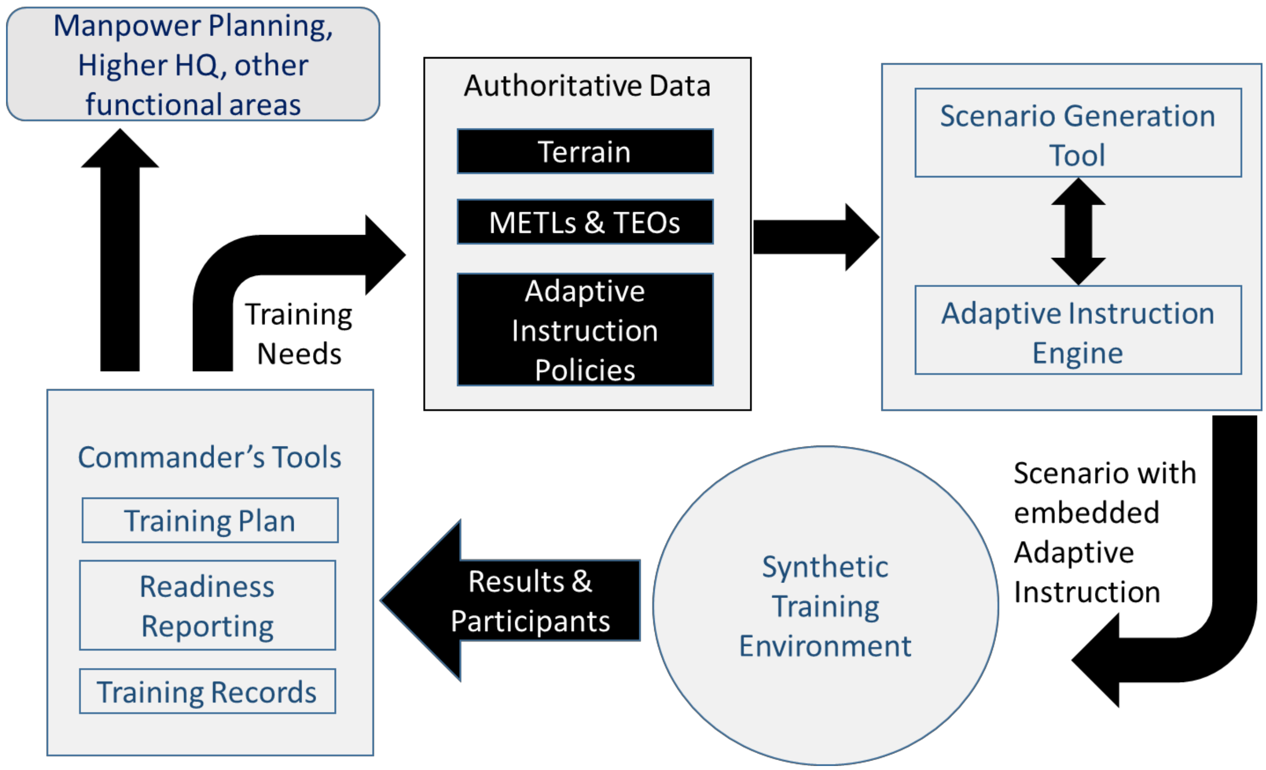


Figure 3. Illustration of how interoperable data (black) could be used to generate training scenarios with embedded adaptive instruction for the STE. METL = Mission Essential Task List and TEO = Training evaluation outline.

Beyond its potential wide adoption, another reason to consider NIEM as a standard is that it has active communities, on-line support and training, and an established development lifecycle. NIEM's framework consists of mature, repeatable, reusable processes and tools that align with an organization's enterprise architecture, allowing organizations to move information quickly and effectively without having to re- build systems or abandon legacy infrastructure. Using NIEM, architects can construct information exchange models based on existing business processes, enabling consistency and a common language across the enterprise. In addition, through exchange reuse, development time is reduced. NIEM can link to or wrap around existing information exchange standards without replacing them. Rather than reinventing the wheel, NIEM allows separately maintained standards to work together, simplifying integration. The NIEM website claims that developers with the skills to design and implement an XML data exchange can learn the NIEM approach in a matter of hours, and that free tools can perform most of the conformance checking.

NIEM is not a single comprehensive model for data exchange (Renner, 2014). In NIEM, interoperability is defined at the IES level. All systems implementing a particular IES are interoperable only for that IES. NIEM domains depend on the ability to create consensus among data exchange designers, and this may be difficult to orchestrate unless all participants believe there is a business case to justify the effort. Finally, NIEM is not a mandated solution for the DoD; exceptions will be allowed, and legacy systems won't necessarily be upgraded.

NIEM is only one method to create standards, and/or represent the semantics of domains. For example, another option is the International Defence Enterprise Architecture Specification for exchange (http://www.ideasgroup.org/0home/), which has a valuable property of explicitly representing states—properties of entities that can change over time. Consideration should also be given to the IMS Global collection of open standards and specifications (https://www.imsglobal.org/institutions.html), which are designed specifically for educational information exchange. Formal ontologies, such as Gellish (http://www.gellish.net/), intended to support interoperability should also be considered.

CONCLUSIONS AND RECOMMENDATIONS

The use of NIEM results in machine-readable information exchanges, which can be used both within and across organizations. It does not specify how to transmit or store data; rather it is a standard way of defining the contents of the information exchange. A NIEM domain manages their portion of the NIEM Data Model and works with other NIEM domains to collaboratively identify areas of overlapping interest and potential reuse. NIEM Core is governed jointly by all NIEM domains and includes elements common across many domains. NIEM adoption is supported by training, documentation, and technical assistance. The purpose of this paper has been to raise awareness of NIEM among learning scientists and engineers, not necessarily to advocate for it. It is recommended that a thorough survey of available existing frameworks, ontologies, and schemas be investigated before committing to any particular one. One potential factor favoring NIEM is that it has been selected as the data exchange standard to be considered first by the DoD.

REFERENCES

Advanced Distributed Learning Initiative (2017). Distributed Learning Gap Report. https://adlnet.gov/.

Brusilovsky, P. (2016). Domain modeling for personalized guidance. In R.A. Sottilare, A.C. Graesser, X. Hu, A.M. Olney, B.D. Nye, and A.M. Sinatra (Eds.). *Design Recommendations for Intelligent Tutoring Systems, Volume 4: Domain Modeling* (pp. 165 – 183). U. S. Army Research Laboratory.

Chief Information Officer for the Department of Defense (2013). Memorandum on Adoption of the National Information Exchange Model within the Department of Defense. http://dodcio.defense.gov/Library/.

Department of Defense (2014). DoD Instruction 8118.01 Mission Partner Environment (MPE) Information Sharing Capability Implementation for the DoD. www.dtic.mil/whs/directives/corres/pdf/811001p.pdf.

Lenat, D. A. (2016). Ontology Alignment of Learner Models and Domain Models. In R.A. Sottilare, A.C. Graesser, X. Hu, A.M. Olney, B.D. Nye, and A.M. Sinatra (Eds.). *Design Recommendations for Intelligent Tutoring Systems, Volume 4: Domain Modeling* (pp. 49 – 55). U. S. Army Research Laboratory.

Office of the Army Chief Information Officer/G-6 (2016). Army Data Strategy v1.0. Information Architecture Division Army Architecture Integration Center. http://ciog6.army.mil/Portals/1/Home/Tabs/Strategy/20160303_Army_Data_Strategy_2016.pdf

Renner, S. (2014). National Information Exchange Model (NIEM): DoD Adoption and Implications for C2. 19th International Command and Control Research and Technology Symposium (ICCRTS). http://www.dodccrp.org/events/19th_iccrts_2014/post_conference/presentations/129.pdf

Robson, R. & Barr, A. (2013). Lowering the barrier to adoption of intelligent tutoring systems through standardization. In R. Sottilare, A. Graesser, X. Hu, & H. Holden (Eds.) *Design Recommendations for Intelligent Tutoring Systems - Volume 1: Learner Modeling* (pp. 5-13). U. S. Army Research Laboratory. https://www.gifttutoring.org/documents/42.

ABOUT THE AUTHORS

Dr. Paula J. Durlach *has conducted and managed adaptive training research, as well as other projects, at U. S. Army and DoD organizations for more than ten years. She is currently at the U.S. Army Research Lab, Simulation and Training Technology Center, and can be contacted at Paula.J.Durlach.civ@mail.mil.*

DISCUSSION COMMENTS

- Author proposed standardizing a domain structure similar to the National Information Exchange Model (NIEM); NIEM provides a standard process for creating information exchange packages (IEPs) and documentation (IEDP); available in various formats; NIEM serves as a “lingua franca” among different knowledge management taxonomies.
- Dr. Sottilare supported Dr. Durlach’s assertion that there might be opportunities to leverage existing standards, interoperability frameworks and/or exchange models like NIEM.

Proceedings of the First Adaptive Instructional Systems (AIS) Standards Workshop

Adaptive Instructional Systems (AISs) are artificially-intelligent, computer-based educational and training systems that tailor instructional content, feedback, and support based on the learning needs of individuals and teams. This workshop was conducted under the auspices of the IEEE Learning Technology Standards Committee and is part of the effort to standardize elements of AISs to expand markets and reduce the skills and time required to produce AISs. Toward this goal, IEEE established a working group under Project 2247.

If you are interested in participating in the AIS standards activity, information about the IEEE P2247 Working Group, news and scheduled events are announced through the IEEE Adaptive Instructional Sciences (AIS) website at: http://sites.ieee.org/sagroups-2247-1/members/. Please sign up today.

About the Editors:

Robert Sottilare, Ph.D. *is the Adaptive Training Research Lead at the US Army Research Laboratory and the IEEE Project 2247 Working Group Chair*

Avron Barr *is the chair of the IEEE Learning Technology Standards Committee*

Arthur Graesser, Ph.D. *is a Professor of psychology at the Institute for Intelligent Systems at the University of Memphis. He is an honorary research scientist in the department of education at Oxford University.*

Xiangen Hu, Ph.D. *is a Professor of psychology, adjunct professor of computer sciences, electrical and computer engineering at the University of Memphis. He is a Professor and Dean in the school of psychology at Central China Normal University.*

Robby Robson, Ph.D. *is a learning technology researcher, innovator, and entrepreneur at Eduworks, Inc. which he co-founded in 2001. He is a past chair of the IEEE Learning Technology Standards Committee (2000–2008).*

Part of the Adaptive Tutoring Series

9 780997 725735