CHAPTER 1 – A Guide to Understanding Learner Models

Arthur Graesser
University of Memphis

Introduction

The core notion of a learner model is not complex intuitively. We simply need to record, represent, and track characteristics of the learner before, during, and after learning. The mission is accomplished when the learner model accommodates all of the variables that are ever considered in the history and future of intelligent tutoring systems, with all variables being adequately represented for all systems. The *theoretical* problem is that the set of variables and their representation is not a closed system, but rather grows over time as a moving target. The *practical* problem is that it is expensive to identify, track, store, update, and later retrieve the ever-growing universal set of variables. The *mapping* problem is that the alignment between the theoretical variables and computer code is often vague, incomplete, or incompatible. The *computational* problem is that complex interactions among learner variables create a combinatorial explosion that is temporally insurmountable. In light of these multiple problems, there is no alternative than to pursue sensible compromises.

The chapters in this section have identified the major challenges in developing a learner model for GIFT. Simply put, there is no consensus in how the community of researchers is to handle the theoretical, practical, mapping, and computational problems. They all offer hints toward a solution, which was their charge. It is difficult to decide whether any of them have answered the challenge. Perhaps yes. Perhaps no.

The chapter by Robson and Barr expresses the need to establish standards. They argue that the standards should be pitched at the macro level rather than the micro level. That is, the community of researchers should agree on the ingredients of major content objects (i.e., knowledge, skills, procedures) but leave it to the individual learning environments to realize the dynamics of mastering these learning objects. This would require an agreement among curriculum experts and system developers in converging on an ideal grain size that differentiates macro and micro.

The chapter by Olney and Cade recommends that researchers take stock of the pedagogical strategies offered by researchers and to identify the learning model variables that support such strategies. The Army Research Lab conducted a systematic study to identify the strategies from hundreds of studies. Olney and Cade identified the classes of learner model variables that would support these strategies. This is a sensible approach to identifying a complete set of learner variables, including cognitive, social, emotion, and motivation dimensions.

The chapter by Lesgold and Graesser raised the persistent problem of transfer. It is comparatively easy to develop a system that efficiently trains learners on the knowledge and skills of a specific learning object, but it is difficult to do so in a way that transfers to a new learning object with related knowledge and skills. Specific is easy, but general is difficult. The authors emphasize that it is absolutely critical to acquire the materials in a general way during training that transfers to a broad range of situations later on. If not, the learning episode is destined to reside in a very narrow corner of the space to be mastered.

The chapter by Pavlik, Brawner, Olney, and Mitrovic provides a serious comprehensive review of the learning models in ITS applications over the decades. They take stock of the learner models in a variety of ITS frameworks, including step-based cognitive tutors, constraint based tutors, knowledge space models, dialogue systems, and trait-based assessments. They point out the value of systems with branching architectures and identify the alternative grain sizes of the branching, as well as the consequences. They also urge the GIFT to build on the major progress that has been made on previous ITS with general architectures and empirical validation.

The four chapters in this section together have suggested a number of factors that need to be considered as GIFT takes on the challenge of learner modeling. The goal of this section introduction is to provide a landscape of relevant dimensions rather than to offer concrete solutions.

Landscape of Variables for Learner Models

The number of learner variables is potentially large and unlimited but a small number of variables is typically tracked in systems that scale up. For example, the variables tracked in educational systems throughout the country rarely go beyond attendance and 1-3 high stakes tests per year. The collection of fine grained measures and learning portfolios are boutique enterprises that interest researchers and teachers of the future. Consequently, any convergence on an intermediate, manageable set of learner variables will need to satisfy numerous political and practical constraints.

The number of learner variables being considered in ITS research has substantially increased over the decades. There are many measures of learning gains that assess changes between pretests and posttests. The efficiency of learning is measured in economic models that consider how much learning occurs per unit time. The mastery of specific knowledge, skills, and strategies is assessed at ever increasing grain sizes. The learners' engagement and persistence can be tracked by recording time on task, reactions to computer requests, non-invasive measures (e.g., eye movements, body posture), and physiological measures. The emotions and motivation of the learner are tracked by algorithms that mine the log files of the person-system interactions. These behavioral measures are arguably more valid than self-report data, such as rating scales that are contaminated by the learners' metacognitive folklore. Measures of personality, leadership, and social responsiveness are also tracked by algorithms that have evolved from the educational data mining community. Contemporary ITS applications routinely collect thousands of measures during a time span of 2 to 20 hours. The grain size of ITS measurement is currently 3 orders of magnitude beyond the data collected in school systems throughout the country.

Farming and Mining the Landscape of Variables

Researchers need to be selective when analyzing the rich log file data that tracks the learner models. The most straightforward approach is to focus on those raw or composite measures that are anticipated theoretically. This top-down approach is the perfect place to start and impresses most reviewers of academic journals and funding agencies. Large data sets can be "farmed" by researchers in a manner that systematically tests and revises theories in the face of empirical data. Data are selected and organized to test major learning theories of the day, well established ITS applications (see Pavlik et al. chapter), educational standards (such as the Common Core or certification on specialty topics), the learning strategies documented by the ARL (see Olney and Cade chapter), an existing repository of learning objects that are shared by the community (see Robson and Barr chapter), and transfer between learning objects, tasks, and subject matters (see Lesgold and Graesser chapter).

Theories are unfortunately limited and frequently not confirmed. Consequently, there is a need for bottom-up methods to discover new learner measures and patterns from the log files. During the last decade, the field has experienced the evolution of the data mining revolution. New categories of learners are revealed by clustering analyses on learners and on tasks, as well as the tracking of individual learner data over time. Longitudinal research designs (which track individuals over a long period of time) are preferred over cross sectional research designs. Sequences of events in the log files are diagnostic of specific psychological attributes. Once these patterns are discovered, they can be tracked automatically and tested further. This approach is expected to lead to the development of more sophisticated learning theories that have a better chance of scaling up.

Representing Learner Models

ITS developers are known to disagree over the representation of knowledge, skills, and strategies in the learner model. This is because the computational models are very different in the cognitive tutors, constraint-based tutors, knowledge space tutors, dialogue-based tutors, scenario-based tutors, and other classes of ITS architectures. The representation of production rules also may differ in ITS applications that adopt production rules. The primitive elements, symbolic expressions, and quantitative parameters differ among production rules systems because the subject matter applications have very different constraints.

It is too early in the history of ITS development to force researchers to adopt a particular representation of knowledge, skills, and strategies. The SCORM initiative never was able to achieve such a lofty objective, even though there were some discussions to try. Researchers are deeply wedded to their pet computational architectures and algorithms as they pursue cycle after cycle of model testing. With this context in mind, GIFT developers may consider focusing on specific exemplar ITS applications that have proven the test of time and empirical validation (see Pavlik et al. chapter). There are a limited set of successful ITS applications so this approach would be a practical first step that is within reach. New ITS applications can build upon these prototypical exemplar ITS applications. The learner models of these exemplars could be expressed in a general formal notation in addition to the actual software residing in a repository of concrete applications. The specific ITS developers would of course need to agree to annotate and release the software. Once the collection of ITS exemplars is available, the time would be ripe to identify a first-cut landscape of variables for learner modeling.

The complexity and variations in representations have left us with some significant barriers in scaling up ITS. There are not enough trained personnel to build new applications on new subject matters because of the idiosyncratic features of the ITS representations. An ideal author would have expertise in the subject matter, cognitive science, information sciences, education, ITS pedagogy, human computer interaction, and sometimes computational linguistics. Authoring tools are often created to minimize this barrier, but there have never been sufficient efforts to build high quality authoring tools. There needs to be systematic R&D on authoring tool development that is tested on personnel outside of the camp of the original ITS developers. We need an applied empirical science of authoring tool development that has analogues to research on writing or to design. To what extent are the learner model representations developed with sufficient fidelity, scope, grain-size, and level of abstraction? How much training is needed for new personnel to develop learner models for new applications? What is the time course and costs of developing new learner models?

GIFT in the Short-Term Horizon

As mentioned earlier, the contributors to this section of the book offered different recommendations for developing the learner model component of GIFT. The recommendations addressed challenges and pressure points that need attention in the roadmap ahead. This final section enumerates some actions that might be considered in the short-term horizon.

(1) A prototype has been developed that implements characteristics of GIFT, including the learner model. A systematic analysis could be conducted on the learner model variables in order to assess the extent to which they cover the variables present in the learner models of different classes of mainstream ITS applications (e.g., cognitive tutors, constraint-based tutors, knowledge space tutors, dialogue-based tutors)

and the pedagogical strategies identified by the ADL. A repository of variables could be assembled with definitions and links to alternative ITS applications.

(2) It is widely acknowledged that it is difficult to achieve successful transfer from one domain or subject matter to that of another. The field would benefit from an analysis of training strategies, representation specification, degree of abstraction, grain size, and other characteristics of learner models of ITS applications that have achieved good transfer.

(3) Shareable learning objects (such as SCORM) have been a persistent dream of many communities that want to scale up advanced learning environments. The field would benefit from a review of the successes and failures of these attempts. This includes an analysis of the level of abstraction and types of representations that are likely to be shareable and serve as standards.

(4) Authoring tools have nearly always been difficult to use for ITS as well as other learning architectures. The field could benefit from a review of empirical research that has systematically analyzed how new personnel use such tools as well as the quantity and quality of their products from the standpoint of learner modeling in particular. New empirical studies are needed that are more systematic than anecdotal.

(5) The time and costs of developing an ITS on a new subject matter has frequently been a focus of questions with respect to scaling up the ITS enterprise. The field would benefit from an economic analysis that helps answer these questions. It is important to segregate the initial up-front costs in developing initial ITS applications, incremental costs in developing new ITS applications that piggyback on existing systems, and scale-up costs after an existing system is ready to be used by thousands or millions of students.

(6) Classes of ITS are ideally tailored to different types of learning, such as strategically guided perception, memory for facts, execution of procedures, explanations of events within complex systems, principle-based prediction/forecasting, and removal of chronic misconceptions in mental models. The field would benefit from a typology and possibly a consensus on what learning environments are appropriate for each type of learning.