Domain Models, Student Models, and Assessment Methods: Three Areas in Need of Standards for Adaptive Instruction

Daniel McCoy

Knowledge Net Consulting, St. Petersburg, FL, 33701, USA

dsm1573@gmail.com

Abstract. This paper offers a proposal for an approach to standardizing Adaptive Instructional Systems (AIS) based on my work in aviation and higher education. Standards to AIS offer potential benefits to developers, educators and learners. They include: interoperability between systems; a common understanding of what constitutes successful learning; and an agreement on how to measure outcomes. On the other hand, AIS standards have the potential to stifle innovation (when they are too strict) or become irrelevant or meaningless (when they are too loose). The cost of re-engineering current systems to accommodate standards presents a further barrier to implementation. In order to maximize the benefits of standardization, I propose a focus on three key areas: domain models, student models, and assessment methods. I offer examples of how standards might be implemented in these areas, outlining the challenges and benefits. Graph models, Bayesian Knowledge Tracing and Item Analysis are discussed.

Keywords: Intelligent Tutoring, Adaptive Instruction, Bayesian Networks, Bayesian Knowledge Tracing, Item Analysis, Domain Model, Student Model.

1 Introduction

This paper addresses the following question: Is the road to personalized learning paved by standardization? I would answer an emphatic yes. And by way of explanation, I take the example of Adaptive Instruction- al Systems (AIS). An AIS is a species of computer-aided instruction that offers a personalized learning pathway to each individual learner. An AIS is "adaptive" in the sense that it takes into account the state of the learner in the process of instruction. The learner's state can be understood as cognitive (knowledge state), affective (emotional state), motivational (volitional state), performative (behavioral state), physiological (limbic state), or any combination thereof. An AIS is instructional in that it involves the transference of knowledge, skills, or abilities from a domain model to the mind/body of the learner. And an AIS is a system by means of

Back to Table of Contents

a structured, algorithmic approach to the process of instruction through the interpretation of the learner's state. The extent of the efficacy of the instructional approach is measured by performance outcomes, usually in the shape of assessment activities.

This is a complex arrangement of cascading constructs. In an attempt to simplify this arrangement for the sake of discussion, I will focus only on the cognitive state as it applies to the design and implementation of an AIS. In so doing, I hope to illustrate why standards are necessary to the design of an AIS, as well as to the practical implementation of the personalized approach to instruction. I argue that we need standards so that proprietary algorithms across disparate instructional systems can make meaningful decisions about (operations on) student learning. These standards should provide frameworks for the stable perception of student states, the construct of domain models, and the interpretation of performance outcomes.

2 Intelligent Tutoring Systems

There is a palpable tension in the design of systems that seek to customize learning experiences through processes that are based on impersonal abstractions of knowledge domains, knowledge transference and performance measurement. But I would argue that we should embrace this tension and we should view standardization as the opportunity for a common understanding of complex processes that give rise to emergent phenomena in the cultural sphere. Take the example of language as an analogy. Human communication is built on a foundation of words and rules. We have a common understanding of the meaning of words and the rules that govern their use. And yet we are able to generate utterances that are meaningful and unique in nearly infinite ways. We may disagree on appropriate word use ("literally!"), correct grammar ("a whole 'nother issue"), and irregular word forms ("let the data speak for themselves"). But as long as we can understand one another, these arguments are best left for academics, language mavens and grammar teachers [1].

The trick for standards in adaptive instruction is to establish a common understanding of the rules and structures of systems that allows for these systems to be generative (new and novel approaches to learning, definitions of success, dynamic knowledge models, etc.), while also facilitating the exchange of student data. One possible starting point to identify candidates for standardization would be to examine the work that is being done in Intelligent Tutoring Systems (ITS). The ITS field gained prominence in the 1980s, with roots in the early years of Artificial Intelligence in the 1960s [2]. A typical ITS will consist of four interoperable components: an expert module based on a well-structured domain; a learner module that stores the learner state throughout the instructional process; the instructional module that presents learning activities based on the learner state and the expert module; and the user interface, where the action happens [3].

This ITS architecture provides some common ground for interoperability between AIS platforms by setting minimal acceptable requirements for the following. The first is a common understanding of how to structure a domain. The second is the type of data that should be included in the student model. The third, I would argue, should be standards for assessments of student performance. In my view, the minimum standards for an AIS include: an understanding of what constitutes the domain; what successful knowledge transference to the student model looks like; and an agreement on the means to measure that success.

One fruitful approach to domain modeling could be to borrow from graph theory. An immediate benefit to this approach would be the capability to construct probabilistic graphical models such as Bayesian Networks based on the nodes of the domain. For example the elements of a course could be represented as a directional acyclic graph that models concepts from the abstract to the concrete, with terminal nodes that interface directly with performance measures (e.g. learning objectives, elements in a task analysis). These terminal nodes are the site of inner-loop adaptivity, where learning activities are presented based on the performance of the learner on the previous activity. Any node preceding a terminal node is a candidate site for outer-loop adaptivity, where new material is presented based on the learner's current state (e.g. mastery). In this approach, learning content could connect to any node. In order to be able the make comparisons between specific do- mains (e.g. Physics 1001 at University X and Physics I at College Y), content should be text-mineable (i.e. not hidden from search).

3 Domain Models

There are numerous possible approaches to domain models. At the outset it should be acknowledged that domain models are an abstraction of an abstraction, based on convention and no small amount of arbitrary guesswork. In a basic conception of a domain, the model represents the information space that coheres to the collective implicit and explicit declarative and procedural consciousness of experts. This information space is not fixed and uncontested. It is the result of the dynamic social processes of dialog, debate, and investigation in search of a provisional consensus on expert knowledge as a legitimate construct. Thus to imagine an expert domain is to envision an ontological flux—semantic webs of significance situated in time and space, promulgated in the activity of agents. Accepting this view of knowledge as arbitrary social facts— constructs generated in the flow of human activity—does not require us to adopt the jaundiced view of "truth" as an instrument of power as promoted by post-structuralist thinkers [4]. Rather we should take the view of the engineer, where "good enough" is sufficient for action.

Thus the abstraction should be sufficient, given that we accept the domain we model may contain components that are more or less well defined. The point is to model the domain so that it is "good enough" for the purposes of instruction to the novice. The role of standards in this instance is to mediate what constitutes "good enough" for knowledge engineers in developing and interpreting domain models. Given that we are still in the early days of standards development for adaptive instructional systems, perhaps the most prudent approach would be consider methods that are best applied to well-defined do- mains. We can avoid the added complexity of addressing root disagreements between experts about what constitutes the domain itself, and focus attention on which models make most intuitive sense from the perspective of pedagogy.

4 Model Constructs in Commercial Aviation

Under an Advanced Qualification Program (AQP), the FAA monitors the process as well as the product. Instead of basing curriculums on prescribed generic maneuvers, procedures and knowledge items, AQP curriculums are based on a detailed analysis of the specific job tasks, knowledge, and skill requirements of each duty position for the individual airline. Compared to traditional training programs, the AQP process provides a systematic basis for establishing an audit trail be- tween training requirements and training methodologies [5].

The field of commercial aviation affords the opportunity to explore questions involving modeling a well-defined domain. The Federal Aviation Administration provides guidance for a voluntary training pro- gram for crewmembers and dispatch personnel. The AQP is designed to produce training programs that are based on specific job tasks and the associated declarative and procedural knowledge required to perform those tasks. This systematic and data driven approach allows for the implementation of a curriculum based on individual competencies rather than generic maneuvers or prescribed instructional seat time. The foundation of this approach is the Job Task Analysis (JTA), where the functions of a particular job role are articulated into a hierarchical structure of supporting tasks, sub-tasks, and elements (see Table 1). For pilots in particular, these tasks are highly routinized and therefore constitute a well-defined domain.

Table 1. Example Task Analysis for Pilot Flying (PF) and Pilot Monitoring (PM) for Takeoff function.

Task Number	Task Description
2.0	Takeoff
2.1	Perform Normal Takeoff
2.1.1	Assess Performance and Environmental Factors
2.1.2	Perform Takeoff Roll
2.1.3	Perform Rotation and Liftoff
2.1.3.1	Rotate aircraft at VR to target pitch angle [PF]
2.1.3.2	Observe barometric/ADC altimeter increase [PF]
2.1.3.3	Call out positive rate [PM]
2.1.3.4	Retract Gear [PF, PM]
2.1.3.5	Establish Climb Speed [PF]

While task hierarchies are useful for describing procedural and conceptual components, their structure is not a sufficiently flexible representation of true thought and action. I would argue that a graph model presents a preferable method that represents

Back to Table of Contents

the parallel processing that occurs in human cognition. This graph method opens the door to an ontological approach that allows for an overlapping of relations between the domain model, the student model and performance assessment.

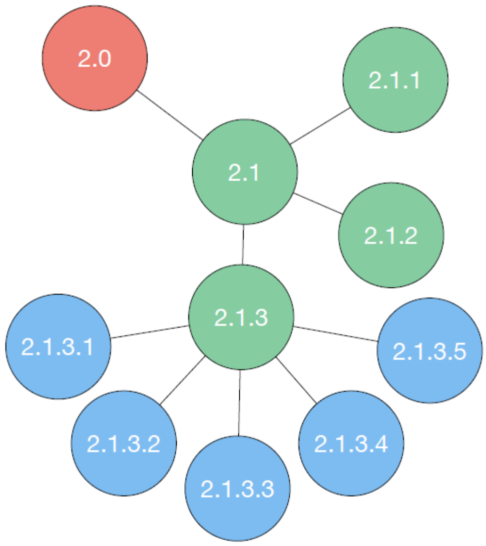

Fig. 1. Phase of Flight JTA example from Table 1 in graph format.

The graph approach not only allows for tasks to be modeled hierarchically—with tasks, sub-tasks and elements in parent-child relations— but illustrating lateral relations as well. For example, several items in the JTA make reference to other related tasks and sub-tasks. The Perform Takeoff (2.0) task makes reference to the following tasks as applicable:

- 9.1 Apply Non-Normal/Emergency System Procedures
- 9.2 Apply Non-Normal/Emergency Operational Procedures
- 10.12 Apply FMS Operation Procedure

Situated in a graph, tasks can start to cluster together, illustrating where skills and knowledge may be transferrable or deployed simultaneously in the flow of activity. Likewise, when ontological domains are considered in relation to one another, overlapping interdependencies can emerge.

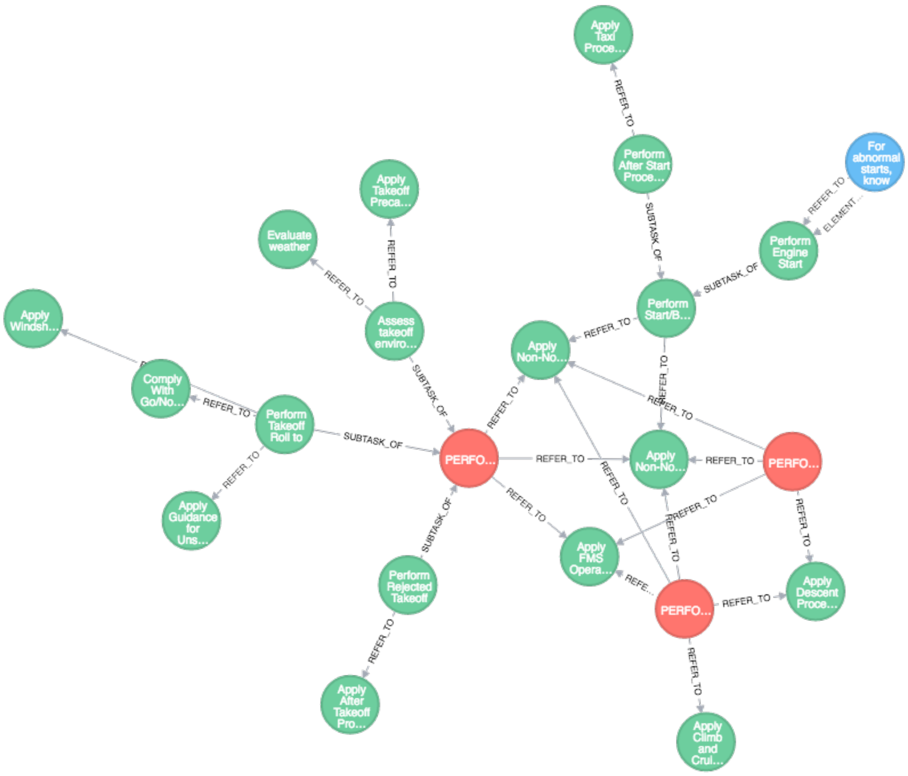

Fig. 2. Connections of related tasks in the Phase of Flight JTA.

Pilot procedural knowledge of the tasks in the phases of flight is de- pendent on the declarative knowledge of aircraft systems, meteorological phenomenon, air-traffic control procedures, and the conventions of radio communications. In commercial aviation, these species of declarative knowledge also constitute well-defined domains. The curriculum is typically mapped in a hierarchical structure similar to the JTA, but with segments, modules and lessons substituting for tasks, sub-tasks and elements. Following the conventions of instructional design techniques, each lesson has associated learning objectives that are measured by assessment activities.

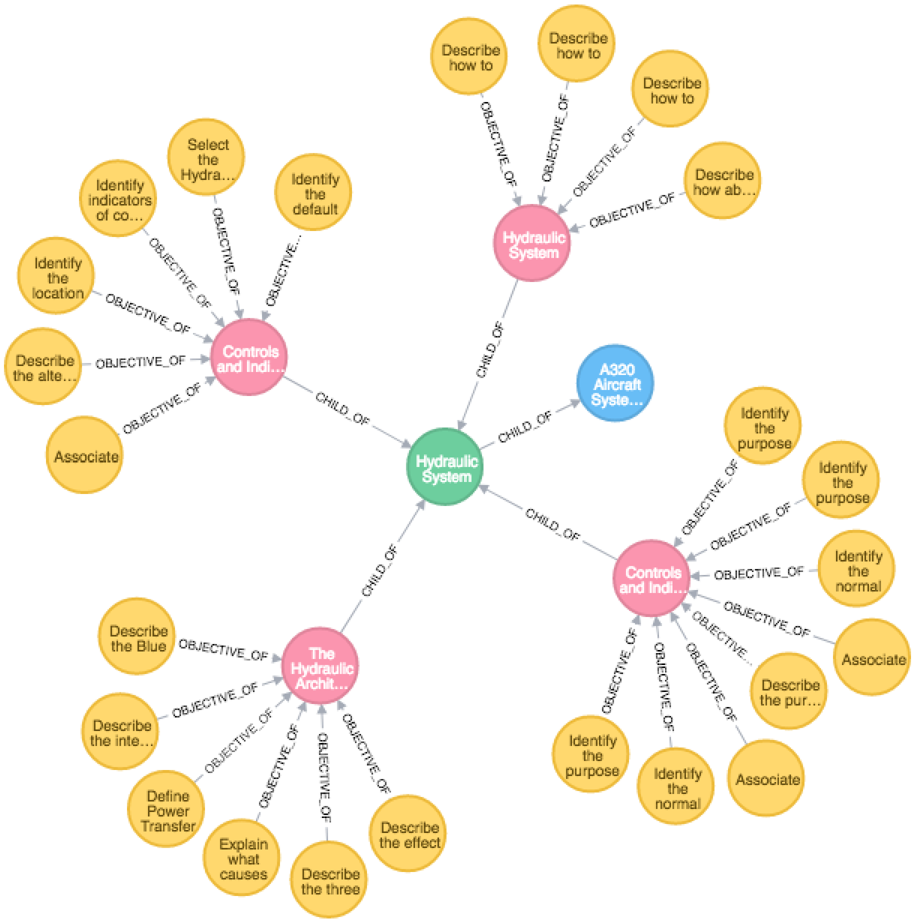

Fig. 3. Domain model example from Aircraft Systems, illustrating the relationships between segments, modules, lessons and objectives.

As an example, the phase of flight designated as “Perform Descent” has associated sub-tasks that requires the monitoring of hydraulic systems. These tasks require the declarative knowledge acquired from the Aircraft Systems domain, where the pilot-in-training learned the controls and indicators for hydraulic system components. This lesson with its associated objectives and activities can be mapped directly to the task of monitoring the requisite systems.

5 Measures of Learning: Application to the Student Model

With the knowledge domain thus structured and mapped according to a network of relations between procedural and declarative knowledge, it is then possible to consider the student domain and the probabilities that the required learning has occurred in order for any given task to be executed properly. In so doing we should consider the probability of mastery of a given node in the network, as well as the impact of mastery on proximal and related nodes. (E.g. does the mastery of hydraulic systems also indicate a high probability of mastery of associated landing procedures?) And finally, we must consider what we mean by "mastery" and understand the measures and procedures by which we arrive to our conclusions.

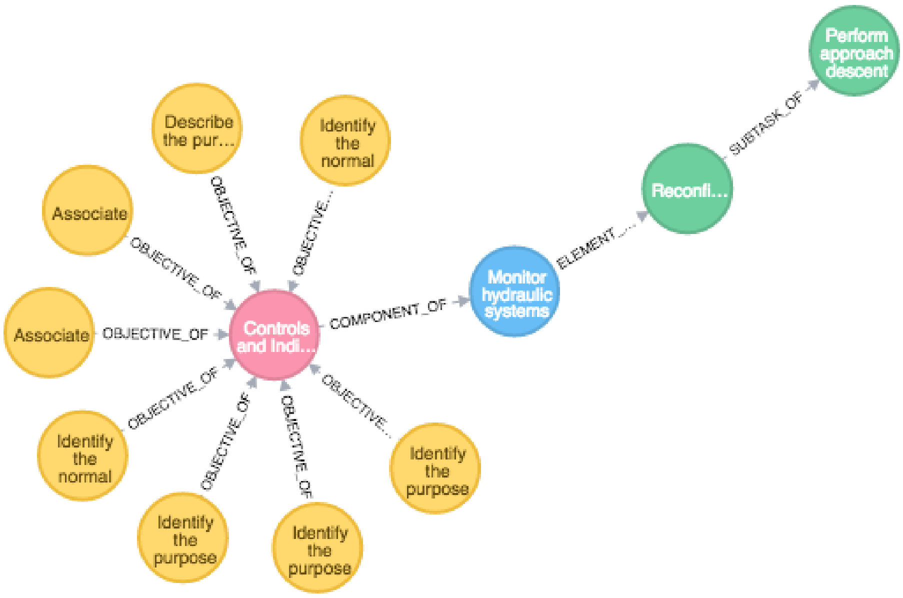

Fig. 4. The connections of the procedural knowledge of the Phase of Flight ontology to the declarative knowledge of Aircraft Systems ontology.

I would suggest three possible techniques to address these concerns: Bayesian network analysis, Bayesian Knowledge Tracing, and item analysis. The first technique gives us the probabilities associated with related nodes, and whether knowledge of one might reliable translate as the likely knowledge of others. The second technique gives us the tools to decide if actual learning is occurring as the student progresses through an AIS. The third provides the tools to differentiate the difficulty levels of various assessment activities.

Back to Table of Contents

5.1 Bayesian Networks

Bayesian networks are probabilistic graphical models, and are typically used to model cause and effect relationships. They are a widely accepted technique for incorporating expert knowledge along with data, de- signed to explicitly represent conditional independence among random variables of interest. The variables of a Bayesian network are situated as nodes in directed acyclic graph with links that represent direct dependencies among these variables [6]. Given a well- structured knowledge domain with a overlapping assessment ontology, it is possible to model the probabilities that particular knowledge components have been mastered based on the performance of learners in related nodes.

5.2 Bayesian Knowledge Tracing

Bayesian Knowledge Tracing (BKT) is a Bayesian Network that mod- els student knowledge as a set of binary variables – one per skill (the skill is either mastered by the student or not). Observations in BKT are also binary: a student gets a problem either right or wrong [7]. The classic formulation of BKT is typically specified by four parameters:

- P(Ln): the probability that a student has mastered a skill prior to solving exercise n;
- P(T): the transition probability from the not-mastered to mastered state;
- P(G): the probability of correctly guessing the answer before skill mastery; and
- P(S): the probability of 'slipping' and incorrectly answering even though a skill has been mastered

BKT has become a popular tool for adaptive learning suites in the past 20 years, and it updates the probability that a student has learned a skill after every item is answered. This property is essential for a mastery model. However, classic BKT doesn't directly factor question difficulty into its probability calculations [8].

5.3 Item Analysis

The question of difficulty can be derived from the statistical analysis on learner performance on particular assessment items. Item difficulty is defined as the percentage of students who answered a test item correctly. Items with as low percentage of correct answers can be assumed to be more difficult that items with a high percentage of correct answers. These difficulty can give added support to assumptions about subject mastery as the student engages the knowledge domain.

6 Conclusion

The question presented at the outset of this paper was intended to frame the preceding discussion as an exploration of possibilities in developing standardized approaches to personalized learning. Adaptive Instructional Systems based on methodologies developed for Intelligent Tutoring Systems present the greatest promise in this regard. Successful systems in this vein are designed with a structured domain model, a corresponding student model, and robust assessment methods. In order for students and AIS developers to benefit from interoperability, certain minimum standards are required. I proposed some ideas for how these standards might be developed using the foundational concepts of the domain model, student model and assessment. I suggest probabilistic graphical models as a potential starting point, with Bayesian Networks, and Bayesian Knowledge Tracing, and Item Analysis as example methodologies where there is high agreement on reliability and predictive value. These specific approaches perhaps should not be established as requirements of an AIS, but might be seen as an added benefit when developers decide to take such approaches. As with the Advanced Qualification Program in relation to commercial pilot training programs, BN, BKT, and IA could be put forth as accepted standards to AIS development, but not a requirement as such. This leaves the door open to further innovation in personalized learning.

References

1. Pinker, S.: Words and Rules: The Ingredients Of Language. Basic Books, New York. (1999).
2. Wenger, E.: Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge. Morgan Kaufmann Publishers, Los Altos, CA. (1987).
3. Sottilare, R. A., Graesser, A. C., Hu, X., Olney, A. M., Nye, B. D., & Sinatra, A. M. (Eds.):. Design Recommendations for Intelligent Tutoring Systems, Volume 4: Domain Modeling. U.S. Army Research Laboratory, Orlando, FL. (2016).
4. Foucault, M: The History of Sexuality, vol. 1. Penguin, Harmondsworth. (1981).
5. FAA: Advanced Qualification Program Advisory Circular. Federal Aviation Administration, Washington, DC. (2006)
6. Conati, C. Bayesian student modeling. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (eds.), Advances in Intelligent Tutoring Systems, Vol. 308, pp. 281–299. Springer, Heidelberg. (2010).
7. Yudelson, M. V., Koedinger, K. R., & Gordon, G. J.: Individualized bayesian knowledge tracing models. In Lecture Notes in Computer Science (Vol. 7926 LNAI, pp. 171–180). Springer, Heidelberg. (2013).
8. Rogers, M. Mastery Model Research Summary. Unpublished manuscript. (n.d.).

Back to Table of Contents