Operational Learning: Leveraging Mission Data to Optimize Skill Development

Kent C. Halverson, Alan Carlin, Kristy Reynolds, David Perlaza, Evan Oster
Aptima, Inc.
Woburn, MA
{khalverson, acarlin, kreynolds, dperlaza, eoster}@aptima.com

ABSTRACT

During a military career, frequently exercised skills appreciate into expertise, while infrequently exercised skills can decay. Decay can be caused by inattention to the skill, which in turn can be caused by infrequent tracking. Although trainee skill states are systematically measured and monitored during formal training (e.g., school house, Initial Qualification Training (IQT), and Mission Qualification Training (MQT)), once trainees are qualified and assigned to operational missions, assessment is less frequent. Training sustainment programs intended to maintain skill proficiency (e.g., Continuation Training (CT)) only require that tasks be accomplished without systematically measuring, storing, or analyzing skill proficiency data. Thus, the problem this paper addresses is that trainee data is not sufficient to determine the nature and magnitude of the skill decay, making it difficult to know the true skill state of military operators at any given time. Fortunately, military operational databases are filled with information related to missions executed, tasks accomplished, tools/platforms used, etc., and can be a rich source of data from which operator skill states can be inferred. In this work, we describe a suite of machine learning data mining algorithms that operate not only on training data stored in Learning Management Systems (LMSs), but also on operational databases, to make inferences about operator skills states that can be used to personalize learning to ensure that only deficient skills are trained. This innovative approach to leverage operational mission data will allow keen insights into operational learning, or the learning that occurs when formal training ends.

ABOUT THE AUTHORS

Kent C. Halverson, Ph.D., is a Senior Scientist at Aptima, Inc. with experience in performance measurement, organizational analysis, leadership development, training evaluation, and social network analysis.

Alan S. Carlin, Ph.D. is a Senior Research Engineer at Aptima. His interests focus on problems of agent and multi-agent planning and intelligent training systems. These include problems of decision-making under uncertainty, communication between members of a team, meta-reasoning among team members, and multi-agent anytime algorithms, artificial intelligence, and machine learning.

Kristy Reynolds is a Scientist at Aptima, Inc. with experience in the areas of job analysis, knowledge elicitation, training needs analysis, competency modeling, training design and evaluation, and performance assessment.

David Perlaza is a Research Engineer at Aptima, Inc. with experience in mathematics, statistics, data mining, software development, and pedagogy.

Evan Oster is an Associate Scientist at Aptima, Inc. with experience in instructional systems design, scenario development, and adaptive training in virtual reality environments.

Operational Learning: Leveraging Mission Data to Optimize Skill Development

Kent C. Halverson, Alan Carlin, Kristy Reynolds, David Perlaza, Evan Oster

Aptima, Inc.

Woburn, MA

{khalverson, acarlin, kreynolds, dperlaza, eoster}@aptima.com

INTRODUCTION

As operational environments become increasingly complex, it is ever more critical that military personnel are trained to, and maintain, proficiency to operate in such environments. Addressing this need is especially challenging as training budgets contract. Most traditional training approaches are manpower intensive, and therefore, an increase in training requires a commensurate increase in costly manpower. However, adaptive training technology exists that can reduce instructor workload and potentially mitigate manpower increases. Adaptive training uses a data-driven approach and leverages machine learning algorithms that can tap into existing data to estimate trainee skill state and make personalized training recommendations.

Truly personalized instruction requires an understanding of what a trainee needs to know or needs to be able to do, an understanding of what they currently know or what they are able to do, as well as what knowledge and skills the training is to develop. Data necessary to develop this understanding is difficult to come by. While some valuable training data exist within current learning management systems (LMS), those systems may lack information with the granularity needed to accurately determine skill levels. For example, traditional training programs typically assess trainees on a pass/fail basis, with only a singular, dichotomous score recorded in the LMS. Although training programs usually identify specific learning objectives or desired knowledge and skill proficiency levels, assessment data are seldom recorded at those more granular levels. In addition, trainee performance data for non-proficient skills seldom get entered in the LMS, with instructors preferring to allow the trainee more time, or attempts, to bring their skill levels up to acceptable proficiency levels, at which point instructors records the proficient assessment scores. This common practice results in trainee performance data with limited variance, since only passing grades are recorded, and therefore, such data is of little value for monitoring trainee skill states over time.

For some military occupational specialties (MOS), data that would aid in training and mitigate some of the data concerns in training LMSs are available in operational systems that store mission data. For example, the intelligence analyst community produces endless streams of analytical products, each of which is a manifestation of an analyst's knowledge and skills applied to a problem. These operational data reflecting mission performance can provide keen insights about skill proficiency and training effectiveness. Training environments are inherently constrained, allowing instructors to easily create and control the conditions and context in which a trainee can exercise specific skills and demonstrate proficiency. Operational missions, on the other hand, are more dynamic and complex, requiring that operators perform skills in real-world settings and conditions. Performance data from operational missions, then, are likely more valid than data obtained in controlled and constrained training exercises. However, extracting those data from disparate operational databases and finding meaningful information is not a trivial task.

In the absence of access to operational performance data, training developers and instructors are limited to costly, time-consuming manual processes to develop training curricula, identify student needs, and deliver training content. Oftentimes, a one-size-fits-all instructional strategy is employed, delivering the same content to learners with little to no consideration for their current level of competence or other individual differences. Specifically, traditional approaches result in individuals receiving more instruction than is necessary based on their competence in certain areas, while also receiving insufficient instruction in areas that are in most need of improvement. Without access to data related to individual trainee proficiency, developing personalized training is challenging.

Exacerbating the issue, current military learning management systems (LMSs) typically use "position" to drive training requirements, though individuals—even with the same position—may vary across several dimensions: (1) the missions that they support; (2) the tools that they use to support such missions; (3) requisite skills to support such missions; (4) the tasks that they conduct on the job each day; and, (5) individual differences such as proficiency. LMS databases and other operational databases contain data relevant to these dimensions, and yet they are not being leveraged to identify individual differences, assess competence, and understand how to create personalized training. These challenges emphasize the need to efficiently and effectively mine both training and operational data that, in turn, allow personalized training to effectively develop competence with limited resources.

To support such personalization, a system is needed that taps training and operational databases to reveal meaningful learner state data (e.g., current knowledge, skills, abilities (KSAs)) and relationships among training programs or missions, and communicates this information to advanced learning management tools that adapt training content and mode via intelligent recommendations. Such recommendations can optimize learning based on the individual's prior training experiences and current operational competence, while achieving global training goals such as maximizing retention and/or minimizing skill decay. Because current LMS databases do little beyond providing static content, and tracking completion of training events and certifications, an advanced system is necessary which leverages data from operational mission databases, analyses these data, and provides individually tailored and operationally contextualized recommendations for training.

OPERATIONAL LEARNING

Operational learning as defined in this paper is a training approach that considers skills exercised in operational missions as evidence of an individual's current skill state. Information from operational missions would supplement training assessment data to provide a more accurate estimate of skill state (see Figure 1). Similar to approaches used by higher level degree programs that offer college credits for work or life experience, operational learning allows military personnel to "get credit" for skills exercised outside of training environments, after performing operational duties. However, while colleges may offer academic credits for prior work experience at the beginning of an individual's college career upon enrollment, operational learning is intended to continuously leverage work experience as recorded in operational databases to more accurately estimate an individual's skill state across a set of skills.

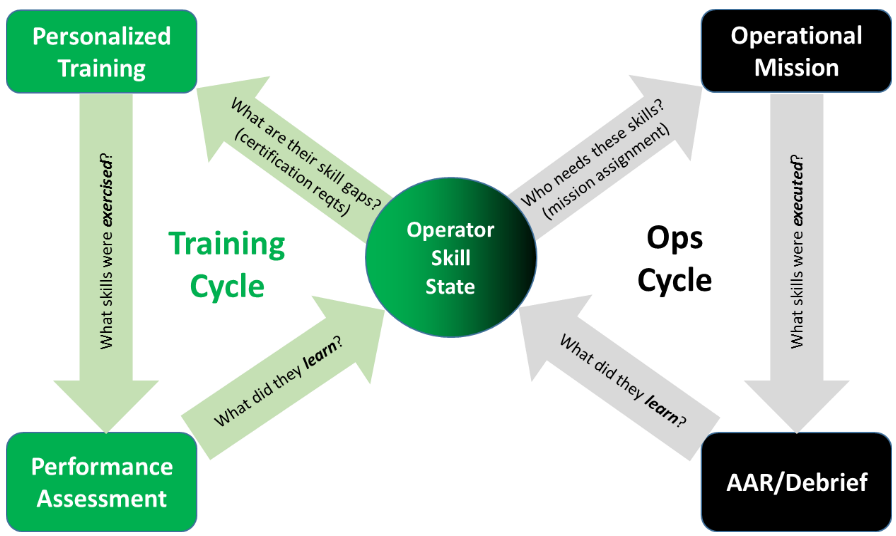

Figure 1. Operational Learning Model

Formal training programs, such as Initial Qualification Training (IQT) and Mission Qualification Training (MQT), are intended to increase skill proficiency to define thresholds required to successfully execute tasks in a mission context. At the conclusion of training, skill proficiency is usually assessed, and in some cases certified via external evaluators. Successful completion of training or certification establishes a baseline skill state for an individual, making them eligible to be assigned to missions. Once an individual leaves the formal training environment, however, their skills are seldom formally assessed until a specific time period has passed and continuation training or re-certification is required, often regardless of whether an individual needs that training or not. This common situation creates a skill assessment blind spot, resulting in those responsible for training to make assumptions about individual skill states, usually erring on the conservative side and assuming that skills have decayed, thus justifying retraining such as Continuation Training (CT).

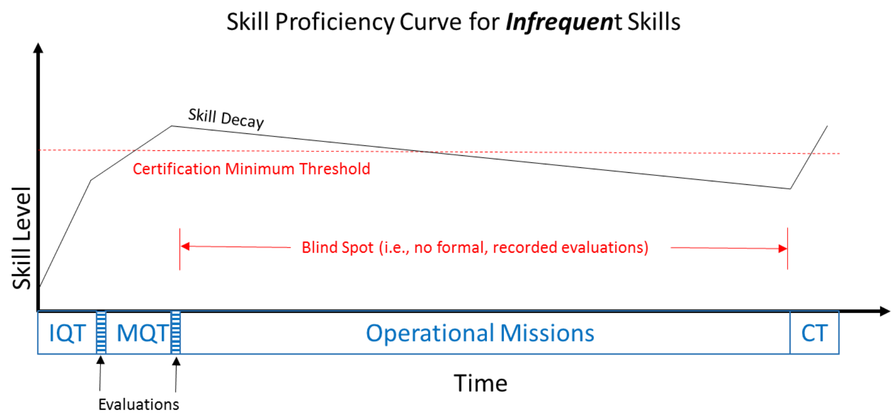

Figure 2. Conceptual Decay Curve for Infrequent Skills

While this conservative approach to training might be efficient for skills infrequently exercised in operational missions (see Figure 2), it is inefficient to require retraining for skills that are exercised frequently and are fully proficient. Continuous monitoring of skill state via analysis of operational data can provide insight about skill decay as well as skill development from proficient to expert level (see Figure 3). As individuals complete operational missions, data about tasks accomplished during those missions accumulate in operational databases. Leveraging these operational databases can reduce the skill assessment blind spot and provide valuable information about an individual's skill state as determined by the tasks completed during operational missions. Adaptive training technology can further leverage these data to more accurately estimate skill state and optimize training recommendations focused on filling skill proficiency gaps without unnecessarily training proficient skills.

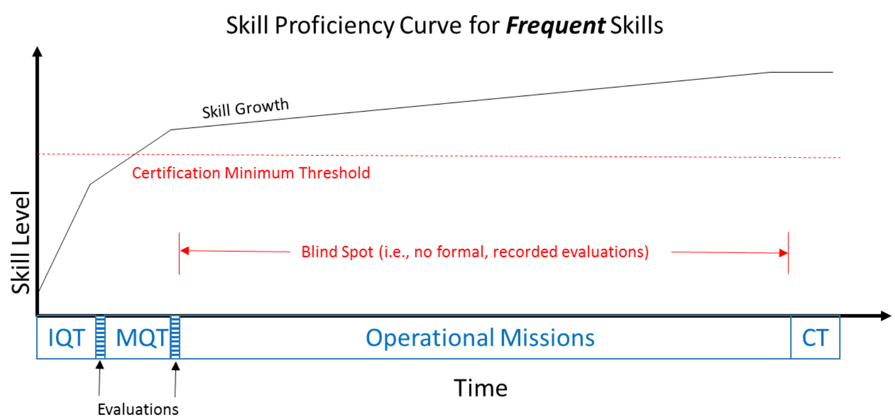

Figure 3. Conceptual Growth Curve for Frequent Skills

ADAPTIVE LEARING

In this section we describe an Advanced Learning Management Platform (ALMP) that uses a validated Partially Observable Markov Decision Process (POMDP) model. Furthermore, it produces adaptive instructional strategies which optimize instructional decisions, given the ALMP model (along with model parameters that customize the model for the individual training domain, an example of a model parameter is the list of skills involved in the training domain). These instructional strategies update their assessments of a trainee in real-time, measure individual trainee performance, and issue recommendations based on their assessment of the trainee's progress; thus the instructional strategies are *adaptive* as training progresses. These model-based instructional strategies have resulted in better learning outcomes than traditional training strategies (Shebilske, Gildea, Freeman, & Levchuk, 2009; Levchuk, Shebilske, & Freeman, 2012). The ALMP model is Bayesian and extensible to other types of data from which skill state can be inferred, such as operational mission databases. Operational mission data can yield two benefits: (1) provide more information about the training domain, and thus can be used to refine and optimize model parameters. (2) provide more information about individual trainee skill state, thus yielding more measures of the individual trainee which, in turn, allows recommendations to be more personalized.

The ALMP model represents the well-validated theory of deliberate practice (Ericsson et al., 1993), which posits that expertise grows best when the learner focuses study and practice on specific, deficient competencies, and receives feedback concerning the effects of this experience. As a Markov model, ALMP models *trainee state* (e.g. "novice", "expert") as a variable which unfolds over a series of discrete time steps, one time step after each training measure or event. As a Bayesian model it contains many variables, and the variable relationships to each other are represented as probabilities, that is each model variable has a probability of being assigned a value, given the value of the related model variables. An example is that the probability that a trainee is in a given state, for a given skill, depends on the trainee's state before the last training event, as well as on the training event itself, as well as the measures of the skill that were part of the training event. Below, we itemize the ALMP model parameters, which are provided as inputs to the modeling software.

Input 1: The set of competencies being trained.
Input 2: Actions—the set of training content available.
Input 3: A model of the Zone of Proximal Development (ZPD; Vygotsky, 1978) that includes a probability that each training event advances each competency, given the learner's current competency levels.
Input 4: A set of probabilities corresponding to measurement error; the chance that a learner is at a given competency level given the measurements taken during training, and
Input 5: A prioritization of training goals.

There are two ways to generate these inputs. First, not the subject of this paper, is for a Subject Matter Expert (SME) to specify them, sometimes with the support of a Scenario Authoring Tool (SAT). The SAT will ask the instructor what the domain competencies are, what training exercises are available, how much each exercise trains each skill, and information about the measures. The second, the subject of this paper, is to infer these model parameters directly from (training and operational) data. This second approach has the benefit of being less costly, and more importantly it can be combined with the first approach to use the expertise of both the instructor and the system. Then, given the model and the parameters, an instructional strategy is built. The instructional strategy can be thought of as software, with the inputs above as well as one additional real-time input, the specific personal measures of the student it is training, as those measures are elicited. From these inputs, the instructional strategy produces the following outputs:

Output 1: An assessment of trainee progress, including assessed trainee competency on each competency in the training domain.
Output 2: A prediction of the effect of all available training content on the current trainee's skill level.
Output 3: An optimal training plan to achieve expertise, beginning with a suggestion to the instructor for the training event that will most likely move the learner towards mastery.

Below, the inputs to the ALMP's POMDP-based model are described in more detail.

Input 1: Competencies

Competencies are domain-specific areas of competence. Different applications of this work have required slightly different terminologies, sometimes we alternatively refer to these as principles or KSE's (Knowledges, Skills, and Experiences). Competency mapping will be further described later in the paper. The model uses a list of competencies, and a number of levels, to determine all possible student states (mathematically, the set of states is the cross-product of sets of all possible levels for all possible skills). An example of a student state is "Intermediate on Skill 1, Novice on Skill 2, and Expert on Skill 3."

Input 2: Actions (Content)

We seek to model decision-making on the part of an instructor. A decision is what action to take. The word "Action" has a specific meaning in POMDP-based model; the set of training actions is the set training options or training content available. In the ALMP, actions are labeled for their applicability for what skills they train, as well as their difficulty with respect to each skill.

The training scenarios that are most likely to enhance adaptive expertise of the given learner mastery of a scenario (S1) are those that lie within the Zone of Proximal Development (ZPD) of the learner. They are neither too easy, too hard, nor near-duplicates of S1. Let |P| be the number of competencies or principles which influence training. A scenario is described as a tuple in $R^{|P|}$. The ith component of the scenario tuple belongs to a range of $1..L$, where L is the number of levels associated with the ith competency. An example of a scenario tuple is $< 1, 3, -1 >$ representing a difficulty of 1 on the first competency, 3 on the second competency, and the third competency is not addressed in the scenario. Scenarios have two effects. The first is to train the student on the principles that are addressed by the scenario. The second is to measure the learner's proficiency on each principle. To augment the ALMP model to leverage operational data, the set of actions is augmented to include the operational missions. Although operational missions were not selected by an instructor, they do have applicability and difficulty with respect to each competency, and so can be used to enhance understanding of student state.

Input 3: Zone of Proximal Development (ZPD)

In any given state of expertise, the learner can be expected to improve on a subset of available training scenarios/content if that training is just difficult enough (not too easy and not too hard) at the targeted competencies. This subset is called the learner's Zone of Proximal Development (Vygotsky, 1978). Scenarios in the ZPD are neither trivially easy nor impossible to solve, but are conquerable with some support from the tutoring system. When a scenario involves multiple competencies, then it should be optimally effective if some or all of its competencies are within the learner's ZPD and the rest have already been mastered by the learner. None of the competencies should be beyond the learner's ZPD. The ZPD is defined by specifying the minimum and maximum threshold of every competency in order to support training.

Input 4: Performance Measurements (Assessments)

Training scenarios also result in a set of measurements, or assessments. Each scenario and competency results in a measurement that estimates the learner state on the competencies exercised by the scenario. For example, item response theory (with certain parameter assignments) says that if a learner competency level is equal to the difficulty of the training item, the learner has a 50% chance of answering correctly. If the learner expertise exceeds the difficulty of the item by a couple of levels, that chance increases to 90+%. Item Response Theory (IRT) is not the only measurement model. For pure vocabulary recall tasks, the accuracy of the measurement may be very high, while for multiple choice identification tasks the probability that a single item correctly measures acquisition of the word may be somewhat lower.

Adaptive training relies heavily on performance assessment methods that generate data at more frequent periods and more granular levels. While human tutors can continually observe and assess student progress to determine optimal training recommendations, adaptive training recommendations depend on trainee performance assessment data that get entered into some type of database, whether manually or automatically. Training simulators can automatically collect and report certain aspects of trainee performance, but much of the training assessment data is manually entered by instructors. As a result, it is critical that instructors enter trainee assessment data accurately and frequently.

Input 5: Prioritization of training goals

This input assigns a reward or value to acquisition of each competency at each level.

Once we have these skills defined, a method of assessing them, and a model that can predict performance given a scenario and learner skill level, we are in a position to reconfigure an ALMP/POMDP agent to tailor instruction accordingly. As discussed earlier, a POMDP requires a notion of competencies or skills that are intended to be improved, along with a model that assess the probability of a certain performance outcome on a scenario. Once initialized with this configuration, a POMDP can be solved to generate a policy that dictates an optimal training plan in order to achieve the desired mastery. This training plan is optimal in the sense that it achieves the training goals stipulated by the designers of the standards in the most efficient way possible, modulo marginal statistical uncertainty.

We also achieve this appraisal through an indirect approach reliant upon IRT. This theory is hinged on the notion of items (and their difficulty), learners (and their skill levels), and observations of learner performance on those items. In this context, we have scenarios that correspond to items in this theoretical framework which learners are required to perform. As defined in the aforementioned standards, these scenarios are designed to elicit certain behaviors and exercise various skills. Thus, successful performance of a scenario entails proficiency in the constituent skills. IRT formalizes this relationship with the notion of applicability, wherein scenarios are applicable to multiple skills to varying extents. Once we have a corpus of historical performance data, and an association of scenarios to skills via the definition of applicabilities, we can fit a model with learner skill and scenario difficulty as parameters in such a way that it has the highest likelihood of predicting the results that were actually observed. The fitted model will produce the desired appraisal of the skills in question.

COMPETENCY MAPPING

One critical aspect of operational learning is translating accomplished mission tasks into evidence of demonstrated competencies, or skills. Recall that adaptive learning technology requires data related to operator skill states to generate a tailored recommendation for each individual. Therefore, it is critical that the data collected from a mission can be linked or mapped to the correct subset of skills. Data related to operational performance can be observed, collected, and recorded in a variety of ways, thus resulting in data residing in databases possessing different structures and parameters. Consequently, data mining algorithms cannot be uniformly applied to different data sets. However, data transformation techniques can be applied to stored data to achieve desired structure and parameters necessary for successful data mining operations.

An important consideration in data transformation is defining standard structures and parameters. The scientific training literature is replete with taxonomies (e.g., KSAs—knowledge, skills, abilities) and the scientific Air Force community has a similar taxonomy (i.e., MECs® —Mission Essential Competencies, which use knowledge, skills, and experiences), but the operational Air Force training community uses a different taxonomy based on JQS (Job Qualification Standard), BOI (Block of Instruction), and TTL (Training Task List). Although arguments could be made about which taxonomy is more valid or effective in this given context, the reality is that the Air Force as an institution has used, and will continue to use for the foreseeable future, the JQS taxonomy, and most training data stored, will possess the JQS structure.

With respect to data parameters, the issue is not which taxonomy to use, but how to aggregate the data at the proper level. The JQSs are essentially phases of a mission consisting of tasks that require a set of skills to successfully perform. Data transformations techniques can aggregate lower-level data to higher-level constructs, but data that exist only in the higher levels (e.g., JQS and BOIs) cannot be similarly decomposed to reveal skills. This latter situation is problematic for advanced learning management techniques since some scores cannot be linked to specific skills that reflect proficiency levels. Therefore, a series of mapping exercises is required to decompose training scenarios and mission tasks to identify critical linkages between skills and performance variables (e.g., outcomes/scores).

As an illustrative example, consider a hypothetical *Communication* skill. Multiple standards make implicit references to various aspects of this skill. For instance, in a TTL, we might find a reference to a task titled, “Coordinating crew work,” which is theoretically observable through chat logs. On the other hand, another task represents a standard that requires demonstration of proper channel communication, which entails, among other things, logging in and using that chat tool. The existence of these similar standards informs the construction of a skill that thematically unifies them, and we draw from chat data found in the multiple source databases to achieve a direct appraisal of a learner’s skill level along this latent attribute.

MAPPING OPERATIONAL DATA TO SKILL PERFORMANCE

Measurement of skill levels is a critical component of the engine that drives adaptive learning. In order to measure these skills and adapt instruction accordingly, it was necessary to define a working set of skills based on existing standards and position-based skill profiles. We mine performance data from various sources and check it against these standards and skill profiles. We use the resulting evaluation to assess the extent to which prior performance has influenced our defined skills.

The algorithms supporting adaptive training systems are inherently data-hungry. As noted previously, the ALMP model can be extended to work with both training data and other types of data from which skill state can be inferred, such as operational mission databases. These databases often contain an enormous amount of data, which can potentially be tapped to refine model parameters and personalize recommendations. However, one limitation to using non-training sources of data, such as operational data, is that performance/proficiency data is often not collected/stored. For example, in the intelligence analysis industry, certain operational databases contain a multitude of “instance” data, such as the number of missions completed in a given timeframe, number of activities performed during those missions, number of platforms worked, total number of targets exploited, and number of products created. This information is useful for assessing general experience and mission experience; however, it does not provide much insight regarding operational performance, skill state, or proficiency level. Although many communities are moving toward collecting more detailed performance/proficiency data during training, which can greatly enhance the effectiveness of adaptive training systems, this granular performance data is still largely missing from operational mission databases.

One way to get around this limitation is to use a combination of other available data (i.e., proxy variables) to infer proficiency, such as the following:

- QA data: Fewer returned products generally indicates higher proficiency.
- Callout confidence: More proficient analysts tend to be more confident in their callouts (mark “probable” more frequently than “possible”).
- Chat rewards: More proficient analysts are likely to accumulate more chat rewards for their contributions or insights.
- Number of chats received: More proficient analysts, with greater knowledge of an AOR, are likely to receive more chats from others.
- Exploitation time: Faster product generation generally indicates higher proficiency, unless the product fails QA due to low quality.
- Chat records: Are there clarifying or follow up questions? More follow-up questions generally indicate lower proficiency.
- Mission/Activity difficulty: Consistently performing missions that are higher in difficulty.

DATA STRUCTURING

Leveraging LMS and operational databases requires a variety of interfaces and queries to pass data and forth and ensure that the data is properly structure to allow data mining algorithms to operate effectively. Considering this, to support an operational learning approach, we must develop a set of data management components and interfaces to facilitate data transfer.

At a high level, the Data Interface component acts as an adapter for the disparate data sources with which we intend to work. The interpreted data output by this component is then consumed by the Data Mining component to produce skill appraisals, item difficulties, and other data needed to initialize the POMDP. The POMDP then uses this information to construct an optimal policy to adapt training. More particularly:

- The Data Interface is implemented as a modular project on the .NET framework, written in C#. It has classes that are suited for interpreting common data persistence schemas, particularly those found on SQL-based RDBMs and CSV flat files. This module interprets the raw data from these sources into a data structure that is interpretable by the Data Mining component. This structure bears the nomenclature and constructs that are idiosyncratic to Item Response Theory.

- The Data Mining component, another project on the .NET framework written in C#, is the implementation of IRT-based algorithms and data structures for assessing skill from performance data. It receives input from the Data Interface via a specific format wherein Items, Learners, and Performance Observations are encapsulated and transmitted. It relies on an algorithm that uses marginal likelihood estimation in a way similar to expectation-maximization to derive item difficulty and learner skill parameters. The estimation of these parameters is optimal in the sense that the model based on these parameters will yield the highest probability of observing the actual performance data that were used as input.

The POMDP, situated in the ALMP platform and implemented in C# on the .NET framework, is initialized with a configuration data structure in a pre-defined XML format that is produced as the output of the Data Mining component. This data structure captures the necessary inputs to initialize a POMDP model.

This data flow from raw sources to the ultimate refinement permits learning adaptation. It can be thought of as an assembly line that sequentially generates more refined intelligence from raw performance data. Once a POMDP agent is configured and instantiated, it runs indefinitely until stopped, listening for inputs from the training system on performance results that were observed and outputting scenario recommendations as a response. This continuous loop is executed over the network, and is characteristic of a typical client-server interaction.

Asynchronously, once we have obtained a measurement of learner skill states, we store this data for use whenever a POMDP agent with an established training policy is commissioned for learning adaptation. This data is stored in the SQL-based SkillState RDBMs referred to previously.

CONCLUSION

Operational learning is an exciting concept that has game changing potential in improving skill development and maintenance for military personnel. This data-based approach will allow more efficient workforce development by tracking skill proficiency longitudinally, providing insights that have not previously been revealed. Operational learning supplements existing training programs, and thus it requires the development of a shared paradigm between the training and operational communities based on common taxonomies and performance assessment methods. Finally, the operational learning paradigm will more tightly couple operational tasks and training exercises, thus allowing additional efficiencies in areas beyond skill assessment.

The operational learning technology described in this paper is currently in development and will result in a prototype that can operate on mock databases, and eventually, real databases. The prototype will be subjected to validation testing during which time algorithms can be iteratively refined, eventually resulting in a stable operational learning tool. The results from the validation tests will be described in a follow-on paper submitted to the 2018 I/ITSEC.

ACKNOWLEDGEMENTS

This work was supported by the Air Force Research Laboratory's 711th Human Performance Wing, Warfighter Readiness Research Division (Contract # FA8650-16-C-6720). The views, opinions, and/or findings contained in this article are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of the Air Force Research Laboratory or the Department of Defense.

REFERENCES

Shebilske. W., Gildea, K., Freeman, J., & Levchuk, G. (2009). Optimizing instructional strategies: A benchmarked experiential system for training (BEST). *Theoretical Issues in Ergonomics Science*, 10(3), 267-278;

Levchuk, G., Shebilske, W., and Freeman, J. (2012). A model-driven instructional strategy: The benchmarked experiential system for training (BEST). In P. J. Durlach & A. M. Lesgold (Eds.) *Adaptive Technologies for Training and Education*. Cambridge, UK: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C., (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406.

Vygotsky, L. (1978). Interaction between learning and development. In Gauvain & Cole (Eds.), *Readings on the Development of Children* (p. 34-40). New York: Scientific American Books.