Presented at: Behavior Representation in Modeling & Simulation Conference (BRIMS). May, 2003.

Variability in Human Behavior Modeling for Military Simulations

Robert E. Wray
Soar Technology, Inc.
3600 Green Court Suite 600
Ann Arbor, MI 48105
734.327.8000
wray@soartech.com

John E. Laird
Electrical Engineering and Computer Science Department
University of Michigan
Ann Arbor, MI 48109
734.647.1761
laird@umich.edu

Keywords:
Modeling architectures, modeling reasoning and decision making, intelligent agents, behavior moderators

ABSTRACT: Capturing and encoding human variability is an increasingly important issue in human behavior models (HBMs) for military simulations. In this paper, we define variability, describe its importance in several military simulation applications, examine the sources of variability in human behavior, and outline an approach to variability that will allow us to explore approaches to low-cost, realistic variability in human behavior models.

Capturing and encoding human variability is an increasingly important issue in human behavior models (HBMs) for military simulations. During the development of TacAir-Soar [1], a detailed model of expert human pilots flying tactical fixed-wing air missions, we faced contradictory requirements with respect to variability. Variability seemed at odds with validation requirements (how could the system be validated if it had any variability at all?) but yet necessary for realistic models (how could we possibly claim to model human behavior without variability?). For TacAir-Soar, we minimized the deliberate introduction of variability, mostly because it was sufficiently challenging to generate correct behavior. Significant variability did emerge in TacAir-Soar from the interaction between a complex environment (other computer and human controlled entities) and the complex knowledge encoded in TacAir-Soar; observers did not feel that the behavior was overly rote or predictable. Thus, we avoided modeling variability explicitly, which would have complicated validation, but realized adequate variability through rich interactions with the environment.

For the last year, however, we have been developing computer-controlled adversaries for building-clearing combat training. In this urban combat ("MOUT") application, realistic variability is a primary requirement because the total behavior space is narrower, the tactics are less prescribed, and trainees must not be able to easily predict and "game" opponent behaviors [2].

We are currently exploring how to create HBMs that exhibit realistic variability in M&S systems, derived from our experiences in developing TacAir-Soar and the VIRTE MOUT adversaries. In this paper, we define variability, examine sources of variability in human behavior, and identify constraints and issues that arise when attempting to model these sources of variability in HBMs. We introduce potential sources of variability in computer-based models of human behavior and describe how they map to the sources of variability in human behavior. We also describe a "strawman" proposal for architecturally supporting realistic variability. We are implementing this proposal in order to explore these questions empirically.

1. Behavior variability

We define behavior variability as differences in observed behavior when entities (human or otherwise) are placed in essentially the same situations. Situations are defined both by the physical environment (such as the terrain, buildings, other entities, and communications with other entities) and the strategic/tactical environment (such as the mission, rules of engagement, and the command structure).

Variability does not imply simple dichotomies such as correct/incorrect or expert/novice. In a given situation, there is often more than one behavior that is consistent with military doctrine, the observed behavior of military personnel, and the expected behavior of adversaries. As illustrated in Figure 1, within the space of all correct or "good" behaviors, different HBMs (or

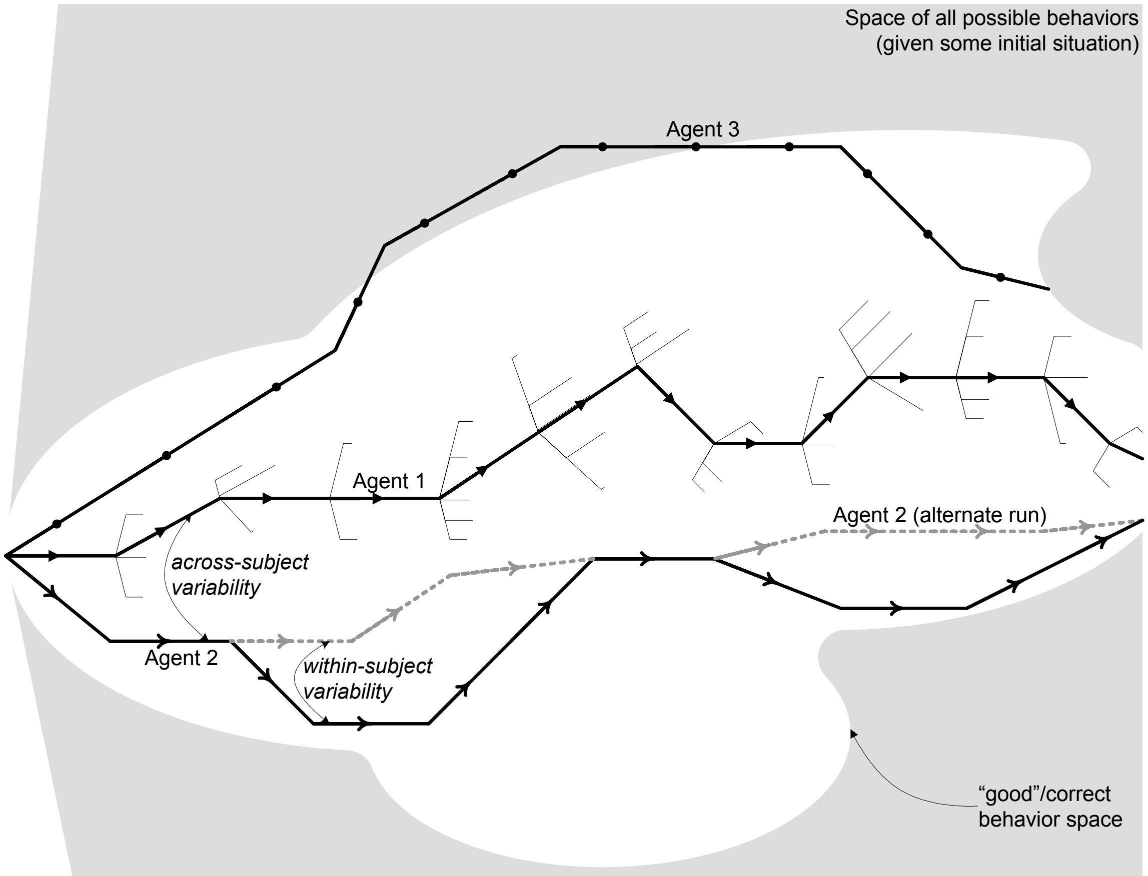

Figure 1 A notional view of behavior space

the same HBM at a different time) can follow different paths through the behavior space. Although the challenge of developing a HBM that is consistent with expected human behavior can be overwhelming in its own right, variability introduces new dimensions for modeling. In most typical HBM systems, models are able to generate only a very small fraction of the total behavior space and concentrate on reproducing models of correct human behavior. Variability requires models that can generate behavior over much larger fractions of the total possible behavior space. This increase includes not only optimal and average behaviors, but also possibly incorrect behaviors.

Introducing variability in HBMs does not eliminate the possibility of generating deterministic behavior. However, to achieve deterministic behavior generation and variability requires that the generation of behavior must be conditional on details of the situation that are not available to the standard observer or participant. For example, an observer might see two situations as the same but resulting in different behavior, but, from the perspectives of the entity that generated the behaviors, the situations could be different. The differences can include both fine-grained details of the current state of the world, not apparent to a high-level observer, and/or differences in the internal state of the entity that would never be observable unless the entity decided to communicate them.

Achieving realistic variability in performance does not in itself require the ability to generate many different options at each decision point. As our experience in TacAir-Soar confirms, small differences in individual decisions can lead to major differences in overall behavior. In Figure 1, the behavior trace of Agent1 includes a representation of the options not chosen at each decision point. Most HBM systems do not model even as many options as are suggested in this notional diagram. However, each time a different decision is made, it changes the situation, possibly leading to different decisions based on the changed situation. In Figure 1, Agents 1, 2 and 3, begin in the same state, but their paths through the behavior space are completely distinct, the result of choosing among different options at the start. Thus, significant differences in behavior can arise from the accumulation of small differences in individual decisions over time. Additionally, the dynamics of physical performance also introduce variability which needs to be captured in HBMs [3, 4].

2. The importance of behavior variability in military simulations

In general, military simulations involve populating simulated battlefields with simulated humans or human-controlled vehicles (tanks, planes, ships, etc), either for training or for the development and evaluation of new weapons systems and tactics. In this section, we discuss why expressing human variability is critical for these applications.

TRAINING

Human behavior modeling is usually used in training applications to populate the battlefield with opponents and teammates. Trainees can directly engage enemy computer forces, fight alongside computer-generated teammates, or command computer forces.

Possibly the most important reason to include variability in HBMs for training is that it prepares the trainee for the variability inherent in human behavior in the real world. For example, when we discussed developing adversary force behaviors for MOUT with SMEs, we suggested creating the "best" adversarial behavior possible – the behavior that would be expected of the most highly trained opponent. This strategy would be comparable to the methodology we used for TacAir-Soar, where we encoded the specific tactics and operational procedures used by military pilots. The SMEs responded that it is critical to expose trainees to the breadth of skill levels in the opponent forces. Untrained forces may behave in ways that are non-optimal and even dangerous for themselves; however the trainees must be prepared to respond to such behavior with appropriate tactics, even if they would not use them against a highly trained opponent. Experts cannot stop and deliberate about an appropriate response; they must know it and execute it immediately.

Variability in computer generated teammates or subordinates is also important because a trainee must learn to work with or command teammates who have different skill levels and learn to recognize and consider variations in subordinates. Attempting to command a heterogeneous fighting force can be very different from commanding a homogeneous one.

Another reason for introducing greater variation in computer forces for training is that without it, trainees may attempt to "game" the situation by taking advantage of the predictability of the computer forces. Learning the characteristics of an opponent is an important skill to learn; however, the predictability of many current computer forces makes it far too easy to execute tactics that are effective against a predictable computer opponent, but would be extremely dangerous in the real world. A final reason for variation is that it can reinforce a trainee's interest and motivation in simulation-based training by challenging the trainee with novel situations.

DEVELOPMENT AND EVALUATION OF NEW WEAPONS AND TACTICS

Just as in training, human behavior models help populate the virtual battlefield when simulation is used for the development and evaluation of new weapons and tactics. The evaluation of the effectiveness of new weapons systems and tactics would be incomplete if there was little or no variability in the computer generated forces (even if we can produce the behavior of average human participants). Testing new systems and tactics against a range of responses to a situation, often at the extremes, allows experimenters to learn not only their strengths, but also their limitations. A new tactic or weapon that works only with the average friendly soldier vs. the average enemy soldier may completely fail when used within the variability that exists across the capabilities and behavior of real forces.

3. Types and sources of variability in human behavior

The types and sources of human variability (see Figure 2) can provide us with guidance and constraint for introducing variability in computer generated forces. A fundamental distinction in variability is whether observed variation refers to the behavior of an individual entity or the behavior among different entities. This section examines these two distinct types of variability and the factors that lead to variability for both types.

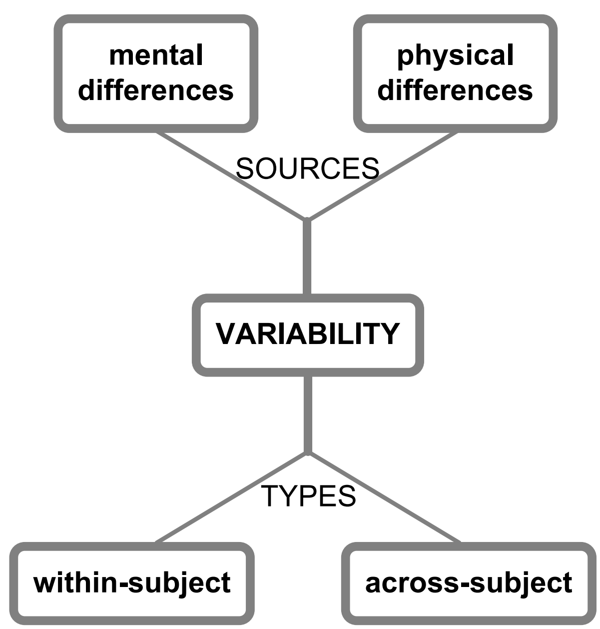

Figure 2 Types and sources of variability in human behavior

ACROSS SUBJECT VARIABILITY

Across-subject variability refers to a situation in which two different people do different things in essentially the same circumstance. For example, an expert soldier may fix on an individual target, shooting until the target is suppressed or disabled, while another, similarly experienced soldier will move from target to target, shooting a few rounds and then focusing on another target without necessarily disabling or suppressing the initial target. Moreover, an experienced soldier may execute a tactic that has been acquired through many hours of training, while a novice may be unable to execute the tactic skillfully or have no knowledge of it at all. Across-subject variability is evident in Figure 1 by the different paths of the three agents in the behavior space.

The fields of Psychology and Human Factors provide a wealth of data for the many reasons variability arises in different subjects. We break down the sources of variability into two classes: physical differences and mental differences (where "mental" is used very broadly). Examples of physical differences include perceptual abilities, level of fitness and health, and physical skill. Examples of mental differences include level of training, education, and intelligence, culture, social standing, religion, personality, memory capacity, and emotional state [5-7].

Many of the "mental" differences can be cast as differences in available knowledge. Consider the factors that influence target selection in urban combat. Henninger, Taylor, et al [8] describe situations in which subject matter experts select and engage different targets in the same situation. Some experts based their decision solely on target proximity; others evaluated the immediate threat posed by potential targets (e.g., returning fire, facing in direction of firer). Others evaluated the longer-term threat of contacts, choosing targets holding weapons with the greatest firepower (e.g., those with semi-automatic rifles, grenades, and rocket-propelled grenades vs. those armed with rifles and pistols). In these cases, the experts drew on knowledge obtained from training, their own experiences, and their assessment of this particular situation to make sometimes different conclusions.

Differences in target selection can also arise from physical differences, such as visual acuity. In the experiments, the SMEs made time-stressed decisions and then reviewed their decisions in an after-action review [8]. One of the subjects reported in this after-action review that he had not seen that one of the targets held an assault weapon and would have chosen it over his proximity-based selection had he recognized the differences in firepower among the potential targets. In this case, a perceptual difference, rather than knowledge differences, led to variation in behavior.

Some mental differences are somewhat difficult to distinguish from knowledge and physical differences, existing at or defining the boundary between "mind" and "body". Differences in behavior arising from mental factors might be difficult for a subject to report, outside of direct cognitive control, yet not simply a physical response. Consider, for instance, the reactions of a well-trained soldier experiencing live enemy fire for the first time in comparison to a seasoned veteran. The emotional responses to this fire are likely to be quite different among the two soldiers and perhaps lead to observable differences in behavior (e.g., keeping a weapon at the ready position) for which both soldiers share similar knowledge. Cultural differences, which are acquired over a lifetime of conditioning and learning, are also important sources of variability for military applications.

WITHIN SUBJECT VARIABILITY

Within subject variability refers to a situation in which an individual takes different actions at different times in effectively the same situation. Perhaps in some cases an expert soldier engages the target in closest proximity while, in others, he engages a target facing his direction, even if it is further away than another.

Once again, behavior differences can arise from either differences in mental or physical state. Within subject differences in mental state can result from a change in the entity's available knowledge through learning. For example, if the first time an entity encounters an enemy platoon and expects them to have only small arms, the entity might choose one tactic for attack. However, after learning that the enemy has more powerful weapons, the entity will probably change its behavior, having learned to improve its evaluation of its best course of action though experience. Other example factors that can influence mental state are differences in the entity's emotional state, level of alertness and motivation.

Differences in physical state can also change behavior, influencing what actions the individual is able to execute, as well as affecting the mental state. For example, a well-trained but heavily fatigued combatant might be observed to shoot much less accurately than under normal conditions. In this case, the variability (in effect on the enemy) is due to the physical effects of fatigue. Of course, it is very difficult to separate the physical effects of fatigue (e.g., how heavy a weapon feels) from the mental effects (motivation to aim carefully, motivation to shoot at all). There is also evidence that fine-grained motor behavior can fall into patterns of variable responses [9].

Of these two types of variability, across-subject variability is potentially more important for the majority of applications of human behavior models. Developing computational models that provide realistic variability will provide a heterogeneous environment in which every entity behaves somewhat differently. No longer would it be the case that developing a tactic that defeats one enemy would defeat them all.

Within subject variability is less important merely because it is unlikely that a human user will have multiple interactions with the same entity. However, for extended interactions with the same entities, such as the adversary forces we are developing for MOUT training, within-subject variability is also important to avoid predictable actions that the trainee can game.

4. Constraints and requirements for modeling variability

4.1 Realistic, individual-level variability

Although there are many advantages to variability in human behavior modeling, there are limitations and concerns that must be addressed. Most importantly, variability for its own sake is rarely desirable. Imagine a behavior system that randomly generated one of its possible actions at each step in its reasoning. These agents would exhibit variability in their own behavior and with respect to one another, and would not be predictable. But they also would lack coherence and salience in their actions with respect to (attributed) goals, and most probably, they would generate behavior inconsistent with human behavior. Using such agents for any purpose could easily be counterproductive. Rather than introduce arbitrary randomness, our goal is to understand the variation that exists in human behavior and to develop and employ mechanisms in agent systems that can produce human-like variability in agent systems.

Although it is unlikely that an application would approach arbitrary randomization as extremely as in this example, the injection of randomness and “noise” in HBMs can lead to similar problems. Realistic variability requires realistic *individual* behavior that must be sustained over the course of a scenario. For example, suppose that, across a population, two tactics appear equally likely in observed behavior. A model that simply randomly chose to use one tactic vs. another for each encounter would reproduce the population mean, but might not reflect the actual behavior of individual humans. It might be that individuals consistently choose one of the two options. This problem is comparable to the problem of creating cognitive models that reproduce population means but not the behavior of any individual subject [10]. Injecting noise into option selection is not the right solution when the goal is to produce individual entity behavior that reflects the behavior of individual humans.

Computational human behavior models should be able to produce both the population-level distributions and accurate individual level models. In addition to the fact that average case behavior may not reflect the actual behavior of any individual, a critical goal of training applications is to provide trainees opportunities to note and take advantage of patterns. In the situation outlined above, the trainee facing a random choice HBM would be missing an opportunity to observe a pattern that will occur in actual combat.

4.2 Balancing variability and validation

Behavior variability has the potential to increase the problems and complexity of behavior validation. The complexity of human behavior models already make them extremely difficult to validate. Introducing variability means that it is not sufficient to expose the model to each situation once and then evaluate it. Potentially, it must be exposed to the same situation multiple times. Determining the required the number of exposures is an open question. Of course, one could argue that a requirement for validation is variation in behavior, so that the validation of a model that does not produce variability is really impossible. We have preliminary work that attempts to address this through automated comparisons of human and computer generated behavior [11].

4.3 Balancing variability and autonomy

A third concern is that variability could make it more difficult to develop scenarios with specific purposes, such as training a specific tactic and exposing the training to specific behavior. The variability in behavior could make it difficult to predict what behavior the behavior model will generate, and thus, what the exact details of the scenario will be. This could most likely be easily controlled by damping the variability, but this exposes an important point about the use of behavior modeling in simulation. Many times, completely autonomous behavior is not required; instead it must be possible to have external input into the behavior of an agent to make sure that the behavior is appropriate for the goals of the scenario. The type of control is analogous to the director of a play that must change the plot dynamically in the face of the actions of non-scripted participants [12].

4.4 Complexity in modeling sources of variability

Variability in human behavior most often arises from complex interactions among the many mental and

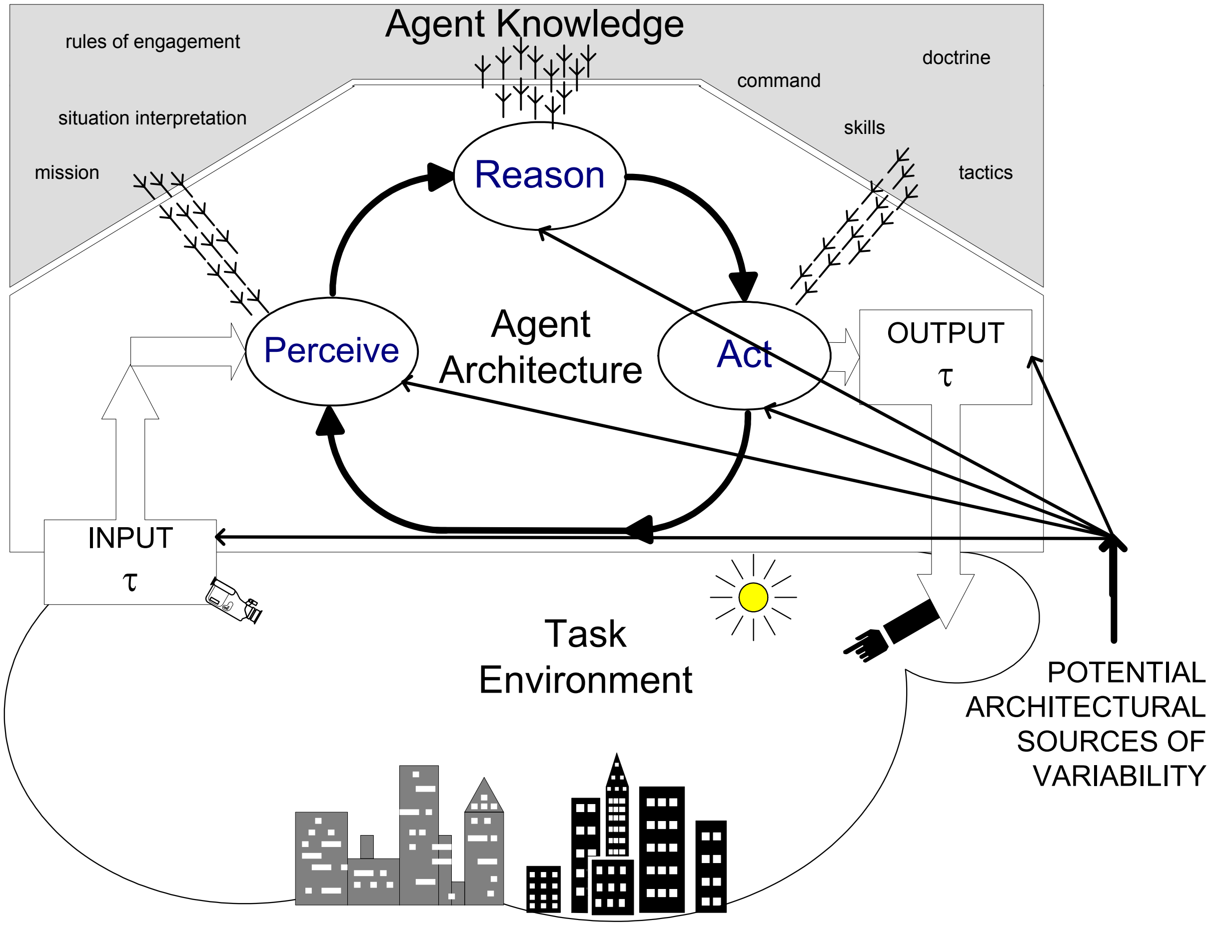

Figure 3 Variability influences all aspects of HBM behavior generation

physical factors identified as sources of variability. What one is thinking of a situation effects emotional state, the effect of which can impact both the physical state and what one is thinking. The subtle and complex interactions of between cognition and just one of these factors is often the subject of research projects in their own right [13-15]. Understanding the influence of these factors on cognition is important research. However, if the goal is to produce realistic variation in individual entity behavior, rather than a computational explanation of some particular phenomenon, it may not be necessary to model the phenomena (and their interactions) explicitly.

5. Computational models of variability

To explore issues in computational models of variability, we introduce a simple model of behavior generation with perception, reasoning, and action components. Figure 3 provides a notional view of such an agent. An agent consists of its fixed processing substrate (agent architecture) and encodings of domain knowledge for particular tasks. Agents receive input from some external task environment. This input must be transformed to a representation that the agent can understand. Internally, the process of perception is mediated by situation interpretation knowledge, allowing the external situation to be interpreted in terms of the current mission, goals, and rules of engagement, etc. Reasoning is the process of generating, evaluating, and selecting goals and actions. Reasoning is influenced by the agent's current situational assessment, its background knowledge, emotional state, and current physical capabilities. The selections of goals lead to actions that are executed in the external environment.

Variability can arise at each stage in this model. Differences in how the situation is perceived, background knowledge, emotional state, and current physical capabilities all can lead to differences in behavior. These differences can arise from either mental processes (perceive, reason, act) or physical interaction components (transfer functions for input and output). For example, differences in situational assessment can be arise from differences in physical ability such as visual acuity, and/or differences in knowledge available for interpretation (an agent might be able to see an enemy weapon but lack knowledge of its capabilities).

The most straightforward path to creating across subject variability in observed behavior would be to create a collection of agents with different knowledge. In Figure 3, knowledge plays a role in perception, reasoning, and action and it seems clear that different knowledge about and interpretations of a situation will lead to differences in human behavior. However, given that the cost of building a single agent is often viewed as prohibitive, the cost of building populations of agents with different knowledge would lead to significant increases in cost, making this option infeasible.

An important factor in this increase in cost is that differences in knowledge leading to differences in behavior must be identified and encoded. Another approach to creating agents with different knowledge would be to allow them to specialize their knowledge through experience and learning. Although variability is not necessarily the goal, this approach is largely the one adopted by researchers in evolutionary computation. However, the learning approach to variability, if not carefully managed, can lead to arbitrary variability. The challenge is to identify and reinforce learning that leads to human-like variability in behavior.

Increasingly, computational modelers are investigating the role of sub-cognitive mental and physical factors in decision making and behavior (e.g., [13-15]). Such models have been used to show, among other things, that they can increase variability in behavior [16]. Thus, they represent a potentially sound source for improving variability. However, the drawback of these models is that they require mature computational accounts of the phenomenon being modeled, and, *a priori*, there is no guarantee that any individual method will provide a rich source of variability. For example, a large number of integrated, architectural models of sub-cognitive phenomena might be necessary to achieve significant variability.

Rather than modeling knowledge differences that lead to different choices or potential sources of variability, our current goal is to create mechanisms that support variability in decision making within the agent architecture and thus simulate some impacts of behavior moderators in terms of variability. Our hypothesis is that, long-term, it will be less expensive to introduce factors that influence the decision making process that can be generalized over many applications rather than attempting to program (or have an agent learn) knowledge differences. We propose to introduce *variability parameters* that can be used to generate realistic within-subject and across-subject variability, but without having to model the sources of variability explicitly. This hypothesis assumes that there are behavior moderators in humans that lead to variability, even when the knowledge of human participants is (more or less) the same.

6. Architectural support for variability in human behavior models

We have begun to explore mechanisms and techniques at the level of the agent architecture and agent knowledge that will allow agents sharing the same knowledge base to exhibit increased variability with respect to the behavior of other agents. The approach is a “strawman,” a simple implementation that will allow us to understand more completely the requirements and evaluation of variability in HBMs.

6.1 Within-subject variability

Because the MOUTBot is our target application for this work, we will illustrate our approach with examples from the MOUTBot. MOUTBot knowledge derives, as in most HBM systems, from interviews with subject matter experts, but also in this case from our own analysis of the tactics of MOUT, and gaming experience in this domain. The MOUTBots are implemented within the Soar architecture [17]. In Soar, knowledge is represented as a collection of operators that are proposed, compared, and selected based on the current situation. Normally, our knowledge acquisition methodology would lead us to identify the best or good choices for a specific situation and to encode those choices. For example, one might only use a grenade at dynamic tactical junctures, or when faced with overwhelming firepower. This kind of model reflects one expert (or at least highly trained) individual because the knowledge encoding is not focused on deliberately identifying and evaluating options, but rather recognizing that a particular operator is the most appropriate for a particular situation. Using this approach, of course, means that across the agents in the simulation and across multiple runs of the simulation, the agents all exhibit similar behaviors. For example, they don't use grenades until tactically appropriate situations and make few tactically surprising decisions. In reality, soldiers make different choices and sometimes mistakes.

Soar allows knowledge engineers to specify choices that are effectively equivalent. For example, for a room clearing scenario, an agent could simply choose randomly between turning right and turning left when entering a doorway. At first glance, this equivalence mechanism alone appears to offer a potential mechanism for introducing within-subject variability. For example, if the rifle, pistol, and grenade weapon selections were equivalent to each other, then the agent would be much less predictable with respect to weapon choice. However, because Soar simply chooses randomly among any equivalent options, the resulting variability would not be realistic. Few combatants will throw grenades capriciously or use pistols when automatic rifles are available. These decisions are much less likely than choosing to use the rifle.

```
Choose a candidate() …

If all choices are equivalent:
1. Average selection values for each
   candidate
2.     If candidate lacks a value,
       add default value
3. Choose candidate from normalized
   probability distribution
```

Table 1. New equivalence selection algorithm

In order to support non-uniform selection distributions, we have slightly extended Soar's basic knowledge representation and modified the Soar decision process, as shown in Table 1, to support the new knowledge representation. Knowledge for proposing and selecting operators can now include a numeric value when indicating equivalent choices. We are still evaluating the semantics of this change but, initially, we are treating it as a probability for a particular candidate. When the options available are all equivalent, the values for each option are averaged (line 1) and then a random choice made from the normalized probability distribution of the averaged values (line 3). For compatibility with existing Soar systems, we include a default value for any options that lack the additional value (line 2). The result is that Soar will continue to choose from a uniform probability distribution when all options have no selection values.

Like all Soar knowledge, the rules comprising the selection knowledge for any option are context sensitive. Thus, there may be any number of rules that express values for an option. Initially, we have decided to average these values (line 1). However, there are many potential choices for the representation of values and the function computed by the architecture. For example, selection could be based on sum rather than average of values for a candidate. Or values could be both negative and positive. We will be exploring the trade-offs of these possibilities but for now we assume values are positive integers between 0 and 100.

Another design issue is the persistence of the selection. If the agent reconsidered selections whenever the active selection probabilities change, its behavior might appear "neurotic," in that it might frequently change its previous decisions (e.g., switch weapons to rifle, then to pistol; then to rifle, etc.). For now we have chosen to assume that a decision, once made, will not be re-decided due only to a change in the asserted selection values. Therefore, if the selection probability of the pistol option increases, some additional factor (perhaps switching targets or moving to a new location) will be necessary before the agent would consider changing its current weapon.

The new mechanism requires a broader knowledge base than would be necessary to create realistic behavior. When variability is desired, the knowledge engineer must identify a range of options rather than one. Consider the target selection example. In the MOUTBots, target selection is based on proximity, which is a valid, realistic algorithm for selecting targets. We will now want to encode multiple target selection strategies and define simple probability distributions among these different strategies. In the long-term, agent development may focus on much more comprehensive efforts to describe and codify behaviors across many classes of subjects.

6.2 Across-subject variability

The selection values and equivalence selection mechanism are only a first step in supporting variability. Selection values represent probability distributions for an individual's options, not for the option selections across the population. Another mechanism is needed to introduce variability across instances of the model.

We have not yet implemented this component of our approach but we have developed a preliminary design. Each time an agent is created, the architecture will create some number of fixed but randomly determined values, that agent's "variability profile." These values notionally represent fixed aspects of the agent's mental and physical state. For example, one value might represent "level of training" and another "susceptibility to arousal." At present, we are not making a commitment to the meaning of particular values.

The variability profile will be used in conjunction with the selection values. The computation of the average selection value (line 1 in Table 1) will be modified as a function of the variability profile. Initially, we can structure agent knowledge to compute any modification although, long term, the moderation functions will be moved to the architecture.

The result of this addition will be that, for any agent, the selection values for options will be constant (meaning that an individual agent will perform consistently over the course of a situation) but that the run-time selection values for different agents will be different from one another, although the agents share the same knowledge bases. Thus, individual agents will generate behavior that allows the recognition of any patterns in its behavior, while supporting the possibility of a number of different patterns across the agents.

7. Test and evaluation of computational models of variability

We have examined sources of variation in humans and a simple approach to realizing greater variation in decision making. However, at this point, we can not claim that this design for variability in decision making will provide human-like decision making. This is an empirical question and will require both data on human variability as well as experimentation to determine how it provides/fail to provide human-like variability.

One reason for proposing this simple approach is that it provides us a tool to begin to evaluate computational approaches to variability. Building models that reproduce both individual subject behavior and an aggregate mean is a difficult problem, even for the more empirically oriented models of cognitive science [10]. Implementing the initial variability profile approach will allow us to begin to explore and to address the problems of evaluation and validation of HBMs with variability and to assess how readily we can realize realistic variability without having to encode knowledge differences or develop sophisticated models of non-cognitive moderators.

8. Conclusions

We have argued that variability is a critical issue for human behavior modeling in military simulations and that the means of achieving variability is not as important as the result in these applications. Thus, we are proposing normative rather descriptive models to introduce variability in HBMs. The variability profile approach may not map onto any human process but will allow us to explore how to achieve human-like variability in HBMs, to what extent the total space of human variation can be achieved without introducing distinct agent knowledge bases, and, importantly, how to quantify and evaluate agent variability in comparison to human variability.

A key contribution of this analysis is highlighting the roles of variability in individual and aggregate behavior. Creating models that reproduce the variability of human behavior is not, in itself, sufficient. Rather, HBMs should generate behavior patterns consistent with individual human behavior while also, in the aggregate, reproducing the broad spectrum of human behavior.

9. Acknowledgements

This work was supported by the Office of Naval Research, Virtual Training & Environments Program, contracts N00014-02-C-003 and N00014-02-1-0009. Kelly McCann and K.C. Pflugler provided answers to our many questions about urban combat behaviors. Andrew Nuxoll participated in the development of the MOUTBot.

10. References

[1] Jones, R.M., J.E. Laird, P.E. Nielsen, K.J. Coulter, P.G. Kenny, and F.V. Koss, *Automated Intelligent Pilots for Combat Flight Simulation.* AI Magazine, 1999. **20**(1): p. 27-42.

[2] Wray, R.E., J.E. Laird, A. Nuxoll, and R.M. Jones. *Intelligent Opponents for Virtual Reality Training.* in *Inter-service/Industry Training, Simulation, and Education Conference (I/ITSEC)*. 2002. Orlando, FL.

[3] Newell, K.M. and D.M. Corcas, eds. *Variability and motor control*. 1993, Human Kinetics Publishers: Chicago.

[4] VanOrden, G.C., J.G. Holden, and M.T. Turvey, *Self-Organization of Cognitive Performance.* Journal of Experimental Psychology: General, 2003. **To appear**.

[5] Silverman, B.G., *More Realistic Human Behavior Models for Agents in Virtual Worlds: Emotion, Stress, and Value Ontologies*. 2001, University of Pennsylvania: Philadelphia.

[6] Johns, M. and B.G. Silverman. *How Emotion and Personality Effect the Utility of Alternative Decisions: A Terrorist Target Selection Case Study*. in *10th Conference on Computer Generated Forces and Behavioral Representation*. 2001. Orlando, FL: SISO.

[7] Lovett, M.C., L.M. Reder, and C. Lebiere, *Modeling working memory in a unified architecture: An ACT-R perspective*, in *Models of working memory: Mechanisms of active maintenance and executive control*, P. Shah, Editor. 1999, Cambridge University Press: New York, NY. p. 135-182.

[8] Henninger, A.E., G.E. Taylor, J. Surdu, and C. Jacquet. *Using COTS to Capture Deliberate and Reactive Weapons Firing Behavior: Lessons Learned in Knowledge Acquisition.* in *2003 Conference on Behavioral Representation in Modeling and Simulation (BRIMS)*. 2003. Scottsdale, AZ.

[9] Riley, M.A. and M.T. Turvey, *Variability and determinism in motor behavior.* Journal of Motor Behavior, In press.

[10] Estes, W.K., *Traps in the route to models of memory and decision.* Psychonomic Bulletin and Review, 2002. **9**(1): p. 3-25.

[11] Wallace, S.A. and J.E. Laird. *Toward Automatic Knowledge Validation*. in *Eleventh Conference on Computer Generated Forces and Behavioral Representation*. 2002. Orlando, Florida.
[12] Magerko, B. and J.E. Laird. *Building an Interactive Drama Architecture*. in *First International Conference on Technologies for Interactive Digital Storytelling and Entertainment*. 2003. Darmstadt, Germany.
[13] Gratch, J. and S. Marsella. *Tears and Fears: Modeling emotions and emotional behaviors in synthetic agents*. in *5th International Conference on Autonomous Agents*. 2001. Montreal, Canada.
[14] Chown, E., R.M. Jones, and A.E. Henninger. *An Architecture for Modulating Agent Decision Making Through Emotions*. in *Sixth International Conference on Autonomous Agents*. 2002. Bologna, Italy.
[15] Jones, R.M., K. Neville, and J.E. Laird. *Modeling Pilot Fatigue with a Synthetic Behavior Model*. in *Seventh Conference on Computer Generated Forces and Behavioral Representation*. 1998. Orlando, Florida.
[16] Henninger, A., R.M. Jones, and E. Chown. *Behaviors that Emerge from Emotion and Cognition: Implementation and Evaluation of a Symbolic-Connectionist Architecture*. in *Proceedings of the 2003 Autonomous Agents and Multiagent Systems (AAMAS 2003)*. 2003. Melbourne, Australia.
[17] Laird, J.E., A. Newell, and P.S. Rosenbloom, *Soar: An architecture for general intelligence.* Artificial Intelligence, 1987. **33**(3): p. 1-64.

Author Biographies

ROBERT E. WRAY is a Senior Scientist at Soar Technology. He received a Ph.D. in computer science and engineering from the University of Michigan in 1998. His research and development experience includes work in agent-based systems, machine learning, knowledge representation and ontology. Previously, he was assistant professor of computer science at Middle Tennessee State University.

JOHN E. LAIRD is a Professor of Electrical Engineering and Computer Science at the University of Michigan and Associate Chair of the Computer Science and Engineering Division. He received his Ph.D. in Computer Science from Carnegie Mellon University in 1983. His research interests spring from a desire to understand the nature of the architecture underlying artificial and natural intelligence. He is one of the original developers of the Soar architecture and leads its continued development and evolution. From 1992-1997, he led the development of TacAir-Soar, a real-time expert system that simulates military pilots and is used for training in large-scale distributed simulations. Most recently his research has broaden to include creating human-like AI systems in computer games. He is a Fellow of AAAI and founder of Soar Technology, Inc.