Autonomous Generation of Intelligent Patterns of Life

David V. Pynadath, Ali Jalal-Kamali, Chirag Merchant
Institute for Creative Technologies
University of Southern California
Los Angeles CA 90094 USA
{pynadath, ajalalkamali, merchant}@ict.usc.edu

ABSTRACT

To prepare Soldiers for missions in densely populated environments, simulation-based training needs to immerse them in a similarly populated training environment. One challenge in simulating such an operationally realistic experience is re-creating the same patterns of life that the real population exhibits in response to alternative courses of action. The complexity of such behaviors and of their interdependency with military operations have motivated a need for AI methods that can generate realistically dynamic patterns of life. Unfortunately, most existing behavior-generation approaches rely on manual scripting of patterns of life, with insufficient flexibility to support the "free thinking" that real-world civilian populations exhibit on a daily basis. We have instead applied a multiagent social-simulation framework, PsychSim, for autonomous behavior generation for individuals and groups, across the range of socio-cultural backgrounds relevant to a simulated operating environment. PsychSim provides reusable mechanisms for the cross-cultural decision-making that forms the basis for the patterns of life implemented in this work (e.g., an individual may choose a route to work that avoids the site of a recent firefight, a crowd of civilians may decide to cheer or protest a BLUFOR unit's actions). We use decision-theoretic agents to choose the behavior they think best advances their goals. Unlike typical agent-based social simulation, the agents' behavior will not be determined by manually authored rules. Instead, the agents will form perceptions of their current situation, their current options, and the relative desirability of those options in terms of their expected outcomes. The agents will thus be sufficiently free-thinking to respond in a robust way to whatever situation they find themselves in, regardless of what path the human behaviors or exogenous events have taken the scenario. We illustrate how this underlying foundation can support a variety of relevant patterns of life taken from operationally relevant scenarios.

ABOUT THE AUTHORS

David V. Pynadath is the Director for Social Simulation Research at the USC Institute for Creative Technologies and a Research Assistant Professor in the USC Computer Science Department. He has published papers on social simulation, multiagent systems, teamwork, plan recognition, and adjustable autonomy. He is the co-creator and maintainer of PsychSim, a multiagent social simulation framework that has been used in interactive simulations for teaching urban stabilization operations, cross-cultural negotiation, and avoiding risky behavior. Dr. Pynadath's work on PsychSim is a key component of his long-term research into applying decision-theoretic multiagent methods to models of behavior. He has developed multiagent systems for applications in social simulation, virtual training environments, human-robot interaction, automated personal assistants, and UAV coordination. He has used such systems to create models of human decision-making in scenarios including ethnic conflict, traffic, classroom violence, negotiation, and disaster response.

Ali Jalal-Kamali is research team lead and a PhD candidate at the Social Simulation Research lab at the USC Institute for Creative Technologies. He is a winner of the Best Teaching Assistant Award at the USC Computer Science Department and also Best Student Paper Award at the Joint IFSA/NAFIPS Conference in 2013. His publications range from interval computations to applied statistical methods for data analysis in social simulation along with multi-agent systems and network influence. His prior work has explored the network connections in international relations data with respect to the actors and their actions. His main focus is on behavioral and perception models of people's decision-making process in networks of social influence.

Chirag Merchant is a software engineer based in Los Angeles, CA, USA. He has developed software professionally for 20 years. At USC's Institute for Creative Technologies, he develops prototypes, simulations, game-based training applications, educational games, and research support software. He has led the development of applications used to train leaders, treat PTSD, prevent sexual harassment and assault, deliver survivor testimonies, teach children AI, learn foreign languages, and visualize simulations. He holds a Master's degree in Computer Science from the University of Southern California with a specialization in Multimedia and Creative Technologies.

Autonomous Generation of Intelligent Patterns of Life

David V. Pynadath, Ali Jalal-Kamali, Chirag Merchant
Institute for Creative Technologies
University of Southern California
Los Angeles CA 90094 USA
{pynadath, ajalalkamali, merchant}@ict.usc.edu

INTRODUCTION

The typical operating environment (OE) has become more and more densely populated. To prepare Soldiers for missions in such environments, synthetic training environments need to immerse them in similarly populated training simulations. This requirement translates into a need for virtual role-players that can form a simulated, but behaviorally realistic, civilian population. To train Warfighters in the social aspects of the Human Dimension, the simulated population must react to trainees' decisions in a flexible and dynamic way that prepares them for dealing with the complex social terrain of a real-world OE.

One challenge in simulating such an operationally realistic experience is re-creating the same patterns of life that the real population exhibits in response to alternative courses of action (Schatz et al., 2012). Being sensitive to even the mundane behaviors of the civilian population can mean the difference between success and failure of an operation. For example, deviations from the local patterns of life led analysts to Osama bin Laden's hideout (Preston, 2012). More generally, an analysis of military scenarios found that a smaller OPFOR could overcome a larger BLUFOR through an information campaign that manipulated the civilian population's movement patterns (Cerri et al., 2017). The complexity of such behaviors and of their interdependency with military operations have motivated a need for AI methods that can generate realistically dynamic patterns of life in training simulations (Folsom-Kovarik et al., 2013).

We have leveraged the intelligent agents of a social-simulation framework, PsychSim (Pynadath & Marsella, 2005), for autonomous behavior generation for individuals and groups, across the range of socio-cultural backgrounds relevant to an OE. These agents represent mainly non-combatants populating the training environment, but their behavior is influenced by BLUFOR and OPFOR decisions. Furthermore, the agents' subjective perceptions change due to BLUFOR and OPFOR decisions, leading them to shift allegiances and choose behaviors that help or hinder BLUFOR and OPFOR objectives.

PsychSim provides reusable mechanisms for the cross-cultural decision-making that forms the basis for operationally relevant patterns of life (e.g., an individual may choose a route to work that avoids the site of a recent firefight, a crowd of civilians may decide to cheer or protest a BLUFOR unit's actions). In particular, we use decision-theoretic agents to choose the behavior they think best advances their goals. Unlike typical agent-based social simulation, the agents' behavior is not determined by manually authored rules. Instead, the agents form perceptions of their current situation, their current decision options, and the relative value of those options in terms of the desirability of their expected outcomes. The agents are thus sufficiently free-thinking to respond in a robust way to whatever situation they find themselves in, regardless of what path the human behaviors or exogenous events have taken the scenario.

RELATED WORK

Unfortunately, most existing behavior-generation approaches rely on manual scripting of patterns of life, with insufficient flexibility to support the “free thinking” that real-world civilian populations exhibit on a daily basis. Many approaches embed subpopulation specifications (including descriptions of their patterns of life) within the terrain data, providing a correspondence between geographic and demographic variations (Hershey & McKeown, 2012; Kerbusch et al., 2012). The generation of patterns of life often incorporates path-planning algorithms that give the entities the flexibility of autonomously choosing a step-by-step route to a goal, although the goals themselves are typically rigidly specified (Blank et al., 2011).

The existing pattern-of-life generation mechanisms do not support much variety in the behaviors of the civilian population. For example, they typically lack a belief mechanism that could react to an information operation by an OPFOR, aimed at disrupting their daily movement routine and spurring them to move in a way that interferes with BLUFOR operations (Cerri et al., 2017). The diversity of potential training scenarios requires a simulated population of virtual role-players who can generate behavior that is robustly consistent psychologically, socially, and culturally. Inconsistencies will limit the behavioral realism of the simulated OE and engender a risk of negative training.

AI researchers have developed a growing body of more sophisticated agent-based frameworks that, like PsychSim, provide a richer model of human behavior for simulation purposes. SoarTech’s Advanced Global Intelligence Leadership Environment (AGILE) uses more complex Soar agents (Rosenbloom et al., 1993) within the context of social networks (Taylor et al., 2004). Like PsychSim, AGILE provides end-users a GUI with which to specify the social environment and an explanation facility through which the agents expose their decision-making. However, AGILE’s Soar agents have a foundation in a cognitive theory from the perspective of an individual agent without the architectural support of theory of mind (specifically, a theory of other minds) that PsychSim agents use to inform their social reasoning.

Like PsychSim and AGILE, PMFserv (Silverman et al., 2006; Silverman et al., 2007) also focuses on higher-level, deliberative reasoning. Furthermore, its creators share our desire to interpolate between the impractical, prescriptive approaches in the AI literature and the informed, but not operational, behavior models in the social science literature (Silverman, 2005). The focus of PMFserv differs from ours in its aim to facilitate interoperability among existing descriptive models for different factors in human decision-making (e.g., arousal, hunger, stress). We instead seek to take inspiration from these descriptive models to extend PsychSim's unified framework of theory of mind to increase both the tractability and fidelity of its prescriptive, decision-theoretic foundation.

PsychSim’s underlying decision-theoretic architecture is more closely related to more recent work on Bayesian models of Theory of Mind from the field of cognitive science. Like PsychSim, this work uses probability distributions to represent the uncertain beliefs that people have of each other and Bayesian inference mechanisms to capture their means for updating those beliefs in response to behavior (Goodman et al., 2006). Also like PsychSim, some of this work uses partially observable Markov decision processes (POMDPs) to represent the decisions people make based on their beliefs, with utility functions representing their goals and preferences (Baker et al., 2011). This work has validated such Bayesian models against human behavior in experimental scenarios, encouraging us that PsychSim’s decision-theoretic architecture is the right choice for a general-purpose social simulation tool.

PSYCHSIM

The foundation of our autonomous behavior generation will be ICT’s social-simulation system, PsychSim. PsychSim provides reusable AI technology for generating multiagent systems capable of populating game environments (Pynadath & Marsella, 2005). PsychSim represents individuals and groups as autonomous agents that integrate two AI technologies: recursive models (Gmytrasiewicz & Durfee, 1995) and decision-theoretic reasoning (Kaelbling et al., 1998). Recursive modeling gives agents a Theory of Mind, to form complex attributions about others, incorporate such beliefs into their own behavior, and enrich the explanations provided to the user. In addition, the agents employ boundedly rational, decision-theoretic reasoning to quantitatively assess risk/reward tradeoffs. Thus, these agents represent a decision-making model that generates behavior by reasoning from a declarative representation of their goals and beliefs.

PsychSim was initially developed in conjunction with the Office of the Assistant Secretary of Defense/Special Operations-Low Intensity Conflict (OASD/SOLIC) and the US Army as an exploratory simulation to determine the impact (especially second- and third-order effects) of an influence campaign on a target population (Marsella et al., 2004). PsychSim was then used to create a countrywide model of Sudan (effort undertaken by MITRE) and to study the role of inter-agency trust in achieving port security.

PsychSim has since been used to drive the behavior of non-player characters (NPCs) in a variety of serious games. It was used to build entities to populate the DARPA-funded Tactical Language Training System (TLTS), an interactive narrative environment in which students could practice their language and culture skills in a simulated foreign city (Si et al. 2005). We also used PsychSim's mental models and quantitative decision-theoretic reasoning to model a spectrum of negotiation styles within the ELECT BiLAT training system (Kim et al, 2009) (funded by the US Army RDECOM). In the negotiation phase of the training environment, the trainee negotiates with a PsychSim agent that has its own goals and possible negotiation moves (e.g., making offers to the trainee, requesting concessions from the trainee, and accepting/rejecting the overall agreement).

USC ICT's UrbanSim simulation-based training game (McAlinden et al., 2014), funded by the US Army RDECOM, used PsychSim agents to generate behavioral responses to an urban-stabilization operation. UrbanSim allows a trainee to take on the role of a battalion or brigade commander who is attempting to maintain stability, fight insurgency and crime, reconstruct the civic infrastructure and prepare for transition in a fictional megacity, populated by multiple individuals and groups, across diverse social and cultural groups.

In each of these game environments, the PsychSim agents respond to the dynamic situation by applying decision-theoretic AI algorithms to update their beliefs about the state of the world and to choose behaviors that best achieve their long-term desires. By modifying the parameters of these behavior-generation algorithms, we can generate different "personalities" of NPCs. For example, by extending the horizon of an agent's decision-making, we can have it project the consequences of its actions further into the future, making it more strategic as we do so. By shortening the horizon, we make the agent more reactive and impulsive. By changing goals, we can change their allegiances and engender conflicts and adversarial faceoffs.

APPROACH

Robust agent-based simulation

In most existing approaches to agent-based simulation, each agent typically chooses actions based on a simple set of rules that capture hypotheses underlying the simulation. Because of the simplicity of each individual's rules, it is possible to simulate thousands of agents behaving simultaneously. While such rules provide a generative model of human behavior, they rely on a domain expert anticipating all situations of interest and the appropriate behavior to perform under each. Because there are no underlying mental processes, these rules proceed directly from conditions to decisions. As a result, changing the hypotheses underlying the simulated mental processes typically requires encoding a completely new set of rules. Furthermore, combining rules from one sphere (e.g., an economic rule stating that a person with a job will not protest vs. a political-science rule stating that an underrepresented person will) often leads to conflicting prescriptions that must be resolved either manually or through ad hoc heuristics (e.g., will an underrepresented person with a job protest or not).

PsychSim instead uses existing decision- and game-theoretic algorithms to generate behavioral policies from its base model. These policies take the form of probabilistic decision trees, which have the same efficiency as the more common rule-based agents. However, these policies can be distinguished from the more traditional rules in that they are automatically derived from first principles, rather than manually encoded. The derivation algorithms achieve optimal policies, ensuring that the resulting behavior is consistent with the underlying beliefs and goals of the entities.

Combination of domain knowledge and data-driven modeling

Machine learning provides powerful methods to use data to build models, but current methods do not allow for the integration of both data and expert knowledge. PsychSim provides a decision-theoretic agent framework that supports learning from data (for increasingly high-fidelity agent behavior). However, it also provides a language for authoring by non-technical domain experts. This language has allowed non-technical domain experts to specify variable values (e.g., how much does a given leader value domestic vs. international prestige) and behavioral constraints (e.g., a given leader would never grant fly-over rights to a coalition partner). PsychSim uses a human-understandable set of variables within its agent models that allows for expert knowledge to be encoded within them.

In addition to supporting such expert knowledge, the ARO-funded OpenMind project developed data-driven algorithms that automatically select the PsychSim models that best match from observed human behavior. In one application of these algorithms, we modeled a disaster-responses scenario in collaboration with the USC Center for Risk and Economic Analysis of Terrorism Events (CREATE), a Department of Homeland Security center (Pynadath et al., 2016). Given an initial model from the domain experts, PsychSim's data-driven models were able to identify a more accurate model for capturing the population, as well as additional models that capture notable subpopulations that deviated from the typical population's response.

Validation of training scenarios

PsychSim's quantitative models not only provide richness in their representation and behavior, but also support rich analytical algorithms that verify the agent models with respect to pedagogical goals. As part of an effort to validate the complex megacity models of UrbanSim, we developed an automated process for translating game logs of human players into agents that captured their behavioral tendencies (Wang et al., 2012). We then used those agents in a Monte Carlo simulation to rapidly generate and profile many more possible game paths and outcomes than would be possible with purely human playtesting. This analysis was able to not only confirm the overall correctness of the manually specified UrbanSim models, but also identify some outlier trainee strategies that violated the intended training but still performed well in the simulation.

In fact, PsychSim's ability to support behavior analysis (instead of just behavior generation) in game settings has proven useful to researchers outside of ICT as well. The Naval Postgraduate School conducted an independent analysis of UrbanSim scenarios similarly leveraging PsychSim's ability to rapidly generate "what-if" outcomes (Vogt, 2013). Researchers at UC Santa Cruz used PsychSim to investigate alternate decision-making strategies in game play using Monte Carlo tree search (Sarratt et al., 2014).

Furthermore, the ability of our agents to systematically weigh the long-term tradeoffs of their options can provide both best- and worst-case outcomes that can be used by other components of an integrated training system. For example, the PsychSim agents that simulated the megacity population in UrbanSim computed the optimal trainee actions and provided them to the tutoring component for after-action review (McAlinden et al., 2009). The information that PsychSim agents can provide about alternative courses of action to intelligent tutoring systems to provide additional training benefit beyond the simulation feedback itself. PsychSim agents can also provide worst-case outcomes, which could be potentially provided to OPFOR agents to inform their action selection when countering the decisions by trainees and BLUFOR units. Varying the amount of such behind-the-scenes information passed along to the adversary provides a game-like method of changing the difficulty of achieving mission objectives in the simulation.

Challenges

To adapt PsychSim agents to the complex simulation needs of realistic training scenarios, this project will address two main challenges:

1. Scalability in simulation: One reason why simple, rule-based agents are the typical basis for population models is that they require very little computational power. It is therefore easy to include large numbers of such agents within a simulation. Richer AI mechanisms, like those used by PsychSim agents, may be more expressive, but they also require more computational resources. While such agents have become more scalable over time (e.g., UrbanSim used over a hundred different agents to represent the city's different groups), training scenarios will operate at a much larger scale, in terms of both number and diversity of agents.
2. Scalability in authoring: Another reason for the use of rule-based agents is that they are easy for human experts to author. The individual agents differ in only a handful of parameter values, and sometimes even use identical parameter values and vary only in location. More expressive AI languages like PsychSim are better suited to capturing psycho-socio-cultural variables and processes, but that expressivity also requires more human input in the authoring process. The burden of authoring effort can become unmanageable when directly applied to the number of agents in a target training scenario.

PSYCHSIM MODEL OF CIVILIAN POPULATION
Autonomous Behavior Generation

State Variables

PsychSim's decision-making framework weighs tradeoffs among potentially conflicting goals, subject to dynamically changing perceptions. We implemented an abstract layer to cover the PMESII-PT variables (political, military, economic, social, information, infrastructure, physical environment, and time) on top of PsychSim's representation language to provide a common structure between the agent's decisions and training scenario design. The following variables are attributes of individual noncombatant agents:

- *health*: The level of this agent's wellbeing, intended to be an "objective" measure that influences physical capabilities and outcomes.
- *wealth*: The level of this agent's general financial resources and stability.
- *hunger*: The level of this agent's physical needs.
- *comfort*: The level of this agent's physical comfort, intended to be a "subjective" measure of wellbeing that can influence the agent's behavior, but may not directly influence other variables.

These variables are chosen as a representation of a civilian's overall wellbeing, so as to provide fine-grained incentivization of the behaviors we wish to model. These variables, if aggregated, would contribute to an overall measure of the Economic status of the area.

Individual agents also have subjective perceptions of the BLUFOR and OPFOR, as well as of any other relevant political entities in the OE (e.g., national government, opposition party, international community):

- *attitude*: The level of support the agent has toward a particular political entity
  This variable, when aggregated over individual agents, would contribute to an overall measure of the Political status of the area.

The BLUFOR and OPFOR are also represented as agents, although we do so to capture the *perceptions* that our noncombatant agents have of them; we expect that other AI systems will drive their actual behavior in simulation. The following variables are currently used to capture those perceptions:

- *resources*: The capacity of this entity for action, thus corresponding to a Military measure. However, we do not draw a strong line here between Military and Economic variables, as economic contributions from supporters would increase the military capacity of an OPFOR, for example. A more fine-grained model would separate these two variables and provide an explicit model of how one translates into the other.
- *tactical_info*: The level of functional information the entity has with regard to the OE and its current objectives. This is a very coarse measure of the Information PMESII-PT component. It provides an easy channel for civilian support (or lack thereof) to influence the entity's capability, as described later.

The OE itself is represented as an agent, albeit one with no capacity for action. The following variables capture the Physical Environment of each region of the OE:

- *risk*: The level of physical insecurity of the region.
- *current_occupancy*: The number of civilians currently in the region.

The civilian, BLUFOR, and OPFOR agents all have variables corresponding to their physical location:

- *x*, *y*: The agent's current location on the 2D map.
  This location affects the dynamics of the agent's PMESII-PT variables by determining what region it is in (and thus what state its immediate Physical Environment is in). Furthermore, it affects the Information received, as many observations are local. For example, the agent will be instantly aware of BLUFOR activity in its field of view, but not of OPFOR activity performed in another part of town.

Reward Variables

The agents also have goals of increasing their wellbeing along some PMESII-PT dimensions (most commonly in seeking to improve their economic and social variables):

- *health*: Incentivizing the agent to avoid traveling to regions with high *risk*
- *wealth*: Incentivizing the agent to go to work
- *comfort*: Disincentivizing the agent from staying away from home
- *hunger*: Incentivizing the agent to go eat.

Varying the priorities of these goals generates agents of psychological and sociological background. For example, increasing the weight that an agent assigns to its *wealth* will make it more willing to travel to high-*risk* regions for work. Conversely, increasing an agent's weight for its *comfort* will it make less willing to leave home for work.

It is obviously infeasible for an author to specify these weights for all of the agents in a reasonably sized population. We instead allow the author to specify a distribution (either uniform or Gaussian) to draw from in choosing the weights across a resulting heterogeneous population. This reduces the authors' task from specifying goal priorities (reward weights) for each agent to specifying only two parameters (high/low for uniform distributions, mean/variance for Gaussian) per goal that are then reused for the entire population. This functionality was used in prior work on modeling population responses to hurricanes, where an emphasis was placed on being able to easily reconfigure the model to capture different population characteristics (Pynadath et al., 2022). Moving the authoring effort to the population level has the added benefit of potentially supporting a data-driven specification. In particular, there may be survey or demographic data that can inform the selection of an aggregate distribution for some parameters, whereas getting data at the individual level would be much harder.

Action Choices

For a patterns-of-life simulation, the civilian agents' behaviors consist mainly of their movement decisions:

- *work*: Selects their place of employment as the destination. Increases the agent's *wealth*, but decreases its *comfort* and increases its *hunger*.
- *eat*: Selects a commercial region with available dining options as the destination. Decreases the agent's *hunger*, but also decreases its *wealth*.
- *go home*: Selects the agent's place of residence as its destination.
- *provide_info*: Provides information to the target entity.

  The first three of these actions are mutually exclusive choices, as the agent can have only one destination at a time. Of course, its current destination may be an intermediate waypoint along its path to a final destination (e.g., an agent needs to go to work, but determines that it can first grab a bite to eat before doing so). The fourth action increases the *tactical_info* of the entity (e.g., BLUFOR or OPFOR) receiving the information.

Variable Dynamics

We tag landmarks in the scenario as contributing to an agent's PMESII-PT status. Some of these landmarks will do so only if the agent is present at that landmark (e.g., a place of employment contributes to an agent's economic status only if it actually goes to work). Others will contribute by being in a functional state (e.g., if a firefight occurs, the region's *risk* will increase). These effects are reusable across agents, although with modification based on individual traits (e.g., nature of employment, proximity to different service providers). This reusability avoids the need for scenario authors to specify these effects when modeling a new OE.

We restrict these effects to be piecewise-linear functions (with possibly stochastic elements) (Pynadath & Marsella, 2004). This restriction enables the compact behavior representation discussed in the next section, while still providing enough richness to at least approximate arbitrary functions in the effects of actions on state variables. It is important to note that these functions' main purpose is to adequately incentivize the civilian action choices, and, as such, only a coarse degree of accuracy is required. In fact, purely linear functions are usually sufficient for behavior generation, even if they deviate greatly from the "true" effect. For example, the effect of *eat* on an agent's *hunger* is to cause a 10% decrease (i.e., *hunger'*=0.9**hunger*). Because the agent wants to minimize its *hunger*, such a decrease will allow it to see the benefit of choosing to *eat*. At the same time, the benefit will decrease as the agent's *hunger* decreases, allowing it to see the benefit of choosing other actions when it is not hungry.

Other action effects are more complex, of course. An agent's *health* is affected by all of its action choices, as moving from one region to another changes the level of *risk* that it is exposed to. The function for updating *health* has negative weights for the *risk* of the agent's current region and its current *hunger*, allowing it to foresee negative consequences for traveling to a region with low physical security.

Behavior Generation

Domain-independent algorithms generate decisions on destinations and paths based on current goals and perceptions. For example, when an area is physically secure, the agents' goal of increasing economic status incentivizes them to move toward their place of employment. However, if an area of the physical environment becomes unsafe, agents will be disincentivized to move through that area and will choose an alternate path. When no satisfactory path exists, the agents may stop going to work altogether. The agents arrive at its decision using a domain-independent algorithm whose output depends on the agents' continuously updating perceptions of its environment. In particular, each agent uses Partially Observable Markov Decision Processes (Kaelbling et al., 1998) composed of its state variables, action choices, action effects, and reward functions to generate its behavior. While a detailed description of POMDP algorithms is beyond the scope of this paper, they provide a procedure for computing the reward that an agent can expect to receive for each of its possible action choices. In our simulations, we perform such a computation over a limited *horizon*, limiting the distance into the future that the agent looks when considering the consequences of its actions. Varying the horizon for different agents will produce different behavior, ranging from more myopic to more strategic, even if they share the same beliefs, goals, etc.

POMDPs provide algorithms for generating an overall *policy* that specifies what the agent would do under every possible state of the world. In general, these algorithms quickly become infeasible as the model complexity grows. For our simulation, we leverage our prior assumption of piecewise-linear action effects to compile the first-principles POMDP models into a compact policy that expresses the agents' behavior as a series of conditionals. In particular, the POMDP projections of action effects over multiple steps translate into the multiplication of piecewise-linear decision trees in our representation. The POMDP accumulation of expected reward over those steps translates into the addition of those trees, scaled by the agents' reward functions. Finally, the choice of the action with the highest expected reward is a maximization step that simply introduces an additional branch at the root of our piecewise-linear decision trees.

The end result of this compilation process is a decision tree, where the branches compute linear functions (i.e., weighted sums) of the state variables and the leaves yield the agent's action choice. One civilian agent in our simulations had the following policy:

IF -1**comfort* – 15.75**health* – 13.0875**hunger* + 4**wealth* – 7.875**risk* (at work) > -62
- **THEN IF** -7**health* – 3.15**hunger* + 1**wealth* – 7.875**risk* (at work) + 4.375**risk* (at home)> -72
  - **THEN** *work*
  - **ELSE** *go home*
- **ELSE IF** 1**comfort* + 8.75**health* + 9.9375**hunger* – 5**wealth* + 4.375**risk* (at home) > -10
  - **THEN** *eat*
  - **ELSE** *go home*

Ignoring the coefficients and thresholds for now, the structure shows a combination of the various incentives and disincentives already described. For example, the branches leading to a decision to *work* have negative coefficients for *comfort, health, hunger,* and *risk* at the agent's place of employment, as working becomes less desirable as these variables become high. The exact numbers represent the specific preferences and attributes of this individual agent, with these numbers varying across the heterogeneous set of agents. While one could imagine an author coming up with basic conditional structure, there is a clear gain in having this compilation process automatically compute the coefficients and thresholds over all the agents, without having to be specified by the author.

This decision-tree form of the agent's policy is easily translated into typical scripting languages for simulated entities. In our usage, we implemented a small Python module to translate PsychSim decision-tree policies into C#, to be used to drive non-playing characters in Unity. After exporting the policies into Unity's entity language, we then successfully imported them into the Rapid Integration and Development Environment (RIDE) (Hartholt et al., 2021) for use in future synthetic training environments. As of publication time, we have been able to successfully run 1000 PsychSim agents in this form in a real-time OE simulation.

Behavior Response to Events

The agents' policies essentially perform dynamic decision making by evaluating possible options contingent on their current perceptions. When external actions (whether performed by BLUFOR/OPFOR units, or injected from a Master Scenario Events List (MSEL)) change the environment, they will subsequently change an agent's decisions, in as much as they change its perceptions. For example, agents who know that an insurgent group detonated an IED in a market will perceive that market to have a less safe physical environment (higher *risk*) and will avoid it. We can also model asymmetric information, as agents who do not know that the IED was detonated will not be deterred from going to the market. In short, because the agent continually reassesses its course of action with respect to its goals and continuously updating beliefs, it can change its plans based on dynamic events, while maintaining consistency with its individual identity.

PsychSim provides a mechanism for agents to automatically update their attitudes toward BLUFOR and OPFOR, based on observable behavior. More precisely, agents weigh every action taken by others (including BLUFOR and OPFOR units) against their own goals. If the action thwarts their goals (e.g., a BLUFOR unit damages a water line, increasing the *risk* in the agents' home region), the agents' attitude toward the performer will become more negative. Conversely, if the action advances their goals (e.g., BLUFOR units rebuild infrastructure in the commercial district, reducing *risk* there), their attitude toward the performer will improve. The quantitative value functions computed by PsychSim's decision-theoretic algorithms provide a principled, reusable mechanism for scaling the resulting attitude shift with the severity of the harm/benefit of the other's action.

The PMESII-PT scoping of agent variables allows for integration with MSEL injects. In particular, both the condition and action of a MSEL inject can be written in terms of PMESII-PT variables, which then map in a straightforward way to the same variables within the proposed agents in the synthetic training environment. The effect of the injects will thus be felt by the agents who share the same variables (e.g., perhaps limited by proximity). The agents' behavioral responses will come about naturally using their same base decision-making mechanism, albeit now using their updated perceptions (as a result of the MSEL inject) as input.

CONCLUSION

Our agents use highly efficient AI algorithms within intelligent agents to generate realistic behaviors for simulated non-combatants in a synthetic training environment. The number and diversity of these agents provides a colorful backdrop to the operation and necessary feedback to trainee decisions. The agents autonomously respond in a way that can help train future Warfighters on what to expect from alternative courses of action within socially diverse OEs. By leveraging robust AI decision-making algorithms, the agents make intelligent behavior choices, consistent with their individual identity. Their behavior maintains an individual consistency regardless of what the trainees choose, what any pucksters do, what the simulated BLUFOR and OPFOR units generate, and what exogenous events are injected. Populating the training environment with these agents will better prepare future Warfighters by incorporating, understanding, and considering the social dimension within their decision-making.

ACKNOWLEDGEMENTS

The authors thank Melvin R. Cape and Howard E. Lee of the TRADOC G-2 Modeling and Simulation Directorate for their invaluable guidance on scenarios to target in designing the behavior space of our agents. The authors thank Anthony Palladino from Boston Fusion for informative discussions and sample data for potentially automating the model-construction process. The work described here was sponsored by the U.S. Army Research Laboratory (ARL) under contract number W911NF-14-D-0005. Statements and opinions expressed and content included do not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

Baker, C.L., Saxe, R.R., and Tenenbaum, J.B. (2011). "Bayesian theory of mind: Modeling joint belief-desire attribution." In *Proceedings of the Annual Conference of the Cognitive Science Society*.

Cerri, T., Laster, N., Hernandez, A., Hall, S.B., Sleevi, N.F., & Johnson, A. (2017). "Simulation of Non-Combatant Population Movement in the Battlespace". In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*.

Folsom-Kovarik, J. T., Schatz, S. L., Jones, R. M., Bartlett, K., & Wray, R. E. (2013). "Scalable Models for Patterns of Life." In *Proceedings of the Conference on Innovative Applications of Artificial Intelligence*.

Goodman, N.D., Baker, C.L., Baraff Bonawitz, E., Mansinghka, V.K., Gopnik, A., Wellman, H., Schulz, L., and Tenenbaum, J.B. (2006). "Intuitive theories of mind: A rational approach to false belief." In *Proceedings of the Annual Conference of the Cognitive Science Society*, pp. 1382-1387.

Hartholt, A., McCullough, K., Fast, E., Leeds, A., Mozgai, S., Aris, T., Ustun, V., Gorden, A.S., and McGroarty, C.. (2021). "Rapid Prototyping for Simulation and Training with the Rapid Integration Development Environment

(RIDE).” In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*.

Kaelbling, L.P., Littman, M.L., and Cassandra, A.R. (1998). "Planning and acting in partially observable stochastic domains" *Artificial Intelligence*, 101(1):99-134.

Kerbusch, P. J. M., Smelik, R. M., Smit, S. K., & Kuiper, F. (2012). "Easy Pattern-Of-Life Generation using Physical and Human Terrain." In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, pp. 2405-2415.

Kim, J.M., Hill Jr., R.W., Durlach, P.J., Lane, H.C., Forbell, E., Core, M., Marsella, S., Pynadath, D., Hart, J. (2009). "BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context" *IJAIED: Special Issue on Ill-Defined Domains*, 19(3):289-308.

Marsella, S.C., Pynadath, D.V., and Read. S.J. (2004). "PsychSim: Agent-based modeling of social interactions and influence." In *Proceedings of the International Conference on Cognitive Modeling*, pp. 243-248.

McAlinden, R., Pynadath, D., & Hill, R.W. Jr. (2014). "UrbanSim: using social simulation to train for stability operations". In C. Ehlschlaeger, ed.: *Understanding Megacities with the Reconnaissance, Surveillance, and Intelligence Paradigm*, pp. 90-99.

Preston, S. W. (2012). "CIA and the Rule of Law." Harvard Law School, Cambridge, MA. April 10, 2012. Public address.

Pynadath, D. V., Dilkina, B., Jeong, D. C., John, R. S., Marsella, S. C., Merchant, C., Miller, L. C., and Read, S. J. (2022). “Disaster world: decision-theoretic agents for simulating population responses to hurricanes”. *Computational and Mathematical Organization Theory, Special Issue on Ground Truth: In Silico Social Science*.

Pynadath, D.V. and Marsella, S.C. (2004). "Fitting and compilation of multiagent models through piecewise linear functions." In *Proceedings of the International Conference on Autonomous Agents and Multi Agent Systems*, pp. 1197-1204.

Pynadath, D.V. and Marsella, S.C. (2005). "PsychSim: Modeling theory of mind with decision- theoretic agents." In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1181-1186.

Pynadath, D.V., Rosoff, H., & John, R.S. (2016). "Semi-automated construction of decision- theoretic models of human behavior". In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*.

Sarratt, T., Pynadath, D., and Jhala, A. (2014). "Converging to a Player Model in Monte-Carlo Tree Search" In *Proceedings of the IEEE Conference on Computational Intelligence and Games*.

Si, M., Marsella, S. C., and Pynadath, D. V. (2005). “Thespian: Using multi-agent fitting to craft interactive drama”. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, pp. 21–28.

Silverman, B.G. (2005). "Toward realism in human performance simulation." In James W. Ness, Victoria Tepe, and Darren Ritzer, editors, *The Science and Simulation of Human Performance*, Emerald Group Publishing, Chapter 9, pp. 469-498.

Silverman, B.G., Bharathy, G.K., Johns, M., Eidelson, R.J., Smith, T.E., and Nye, B. (2007). "Socio-cultural games for training and analysis." *IEEE Transactions on System, Man, and Cybernetics*, 37(6): 1113-1130.

Silverman, B.G., Johns, M., Cornwell, J., and O'Brien, K. (2006). "Human behavior models for agents in simulators and games: Part I: Enabling science with PMFserv." *Presence*, 15(2):139-162.

Taylor, G., Frederiksen, R., Vane, R.R., and Waltz, E. (2004). "Agent-based simulation of geo- political conflict." In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 884-891.

Vogt, B. (2013) ."Two Methodologies to Assess UrbanSim Scenarios" In *Proceedings of the Interservice/Jndustry Training, Simulation, and Education Conference*.

Wang, N., Pynadath, D. V., and Marsella, S. C. (2012). "Toward automatic verification of multiagent systems for training simulations". In *Proceedings of the International Conference on Intelligent Tutoring Systems*, pp. 151–161.