AI Challenge Problem: Scalable Models for Patterns of Life

J. T. Folsom-Kovarik, Sae Schatz,
Randolph M. Jones, Kathleen Bartlett, Robert E. Wray

■ *The patterns of life (POL) present a unique challenge for computational modeling because systems must produce both meaningful regularities throughout a large population and also believable behaviors at the level of each individual. We describe how computational POL modeling integrates diverse artificial intelligence research areas and provides interesting challenges in multiple fields.*

Patterns of life (POL) are emergent properties of complex human social systems such as neighborhoods or even cities (Schatz et al. 2012). In such systems, observable regularities manifest from the interactions of individuals' behaviors and social norms. Simultaneously, these patterns of life impose structure on individual decisions. For example, a pattern of rush hour traffic arises from drivers' decisions to commute at a certain time. Knowledge of rush hour influences individuals' departure times.

Modeling POL is not only an academic pursuit. Before the promulgation of germ theory, medical experts were able to explain a deadly cholera outbreak that decimated one London neighborhood when they reconstructed local patterns of life and related affected individuals to a particular public water pump (Snow 1855). In 2011, sociocultural experts were able to locate Osama bin Laden by recognizing ways his household deviated from the typical pattern of life in the surroundings, behaving much more secretively than his suburban neighbors (Preston 2012).

Building on such social and cultural expertise, computational modeling of POL offers potential to continue extending POL's practical application and theoretical study. In industry, POL models might support decisions about marketing, logistics, network security, or building design. Mili-

tary and law-enforcement models could support observational and small-unit interaction training, intrusion detection, emergency planning, or community relations. Computational POL models will also benefit academic researchers by giving them new tools to generate and validate theories about complex sociocultural patterns.

Finally, POL presents an interesting computational modeling challenge for research in many AI fields because of its complex, multilevel interaction. POL modeling differs from standard agent-based modeling in that not only individual behaviors, but also their collective emergent patterns, are targets of study. At the same time, the behavior of any single individual is potentially important.

As a specific example, the virtual observation platform (VOP) is an immersive simulation designed to train military personnel to interpret indicators before a critical incident, such as a terrorist attack, occurs (Schatz et al. 2012). Pre-event indicators are often ambiguous variations in a typical pattern: a car parked in an odd spot, a sparse crowd in a normally busy public square. Trainees practice identifying and reporting regularities of individual behavior, small group interactions, and population gestalt as they watch a simulation play out in real time. They should see subtle variations in interactions between simulated characters and interpret the meaning underlying any differences. To support these needs, the system requires a POL model.

Challenge: Scalable Models of Patterns of Life

Creating computational models of POL, however, is a significant scientific and technical challenge. To meet it, researchers must achieve simultaneous scalability along three key dimensions: population size, individual fidelity, and automatic behavior specification. In the following text, we outline the POL scalability challenges and argue that AI offers methods and tools that have been used successfully in solving problems with features similar to POL modeling.

Population Size

Generating the behavior of many individuals in real time (or faster) remains a technical hurdle for POL. Even small villages contain hundreds of residents; cities have millions. In the VOP example, trainees might need to see hundreds of simulated characters populating a public marketplace in order to understand the normal range of behaviors and interactions in such an environment. At the same time, many more characters need to populate every surrounding business and city street that the trainees can observe. The large population is vital to providing the required fidelity in background activities and creating an appropriate context for identifying anomalies.

In the current state of the art, researchers contem-

plate scaling to millions of entities through parallelism (Aaby, Perumalla, and Seal 2010). However, narrow individual intelligences are required to model large populations today. For example, fluid dynamics efficiently models entity movement in large crowds (Hughes 2002) but is limited to modeling only physical movements and not the meetings, arguments, and interchanges that take place between friends and strangers.

Individual Fidelity

Scalable POL requires greater breadth in intelligence of individuals, including richness of inputs, behavior choices, and decision complexity in order to produce realistic behavior for individual entities. For example, the displayed behaviors that underlie patterns may need to be highly robust so that they can be carried out despite the interference of human users in the simulation. In cases where humans do not directly influence a simulation, intelligence is still required in order to respond to the many other agents that populate the environment.

Unfortunately, efficient approaches like fluid dynamics models do not model individual intents and actions, and therefore do not provide the fine-grained interactions necessary for many POL applications. Abstract models that completely subsume individual decisions in order to display group behaviors lack believability under fine observation of individuals, and if no goal or reason drives an agent, it is impossible to discern a correct meaning of its behaviors. Conversely, executing a fine-grained model of every individual, at all times, is likely to create unacceptable authoring and run-time computation requirements.

As a consequence, a heterogeneous environment of intelligent entities is necessary in POL systems. This introduces additional run-time challenges: for example, blending relatively few highly intelligent agents with less intelligent entities so that distinctions between them are not obvious and dynamically switching individuals between simple and robust controllers, letting the entire population present large-scale patterns and modeling select entities with increased fidelity only when needed to ensure that vital interactions appear. We are taking this approach in investigating efficient POL representations that support planning and reacting for a small city of entities (Jones et al., 2014).

Behavior Specification

POL model synthesis includes authoring, generalization to diverse settings, and run-time adaptation. The complex interaction of components in generating a pattern of life creates tension between human specification (which is costly in time and labor), automated model creation from real-world data (which may be too tied to specific populations and biased by availability of limited data sources), and synthetic
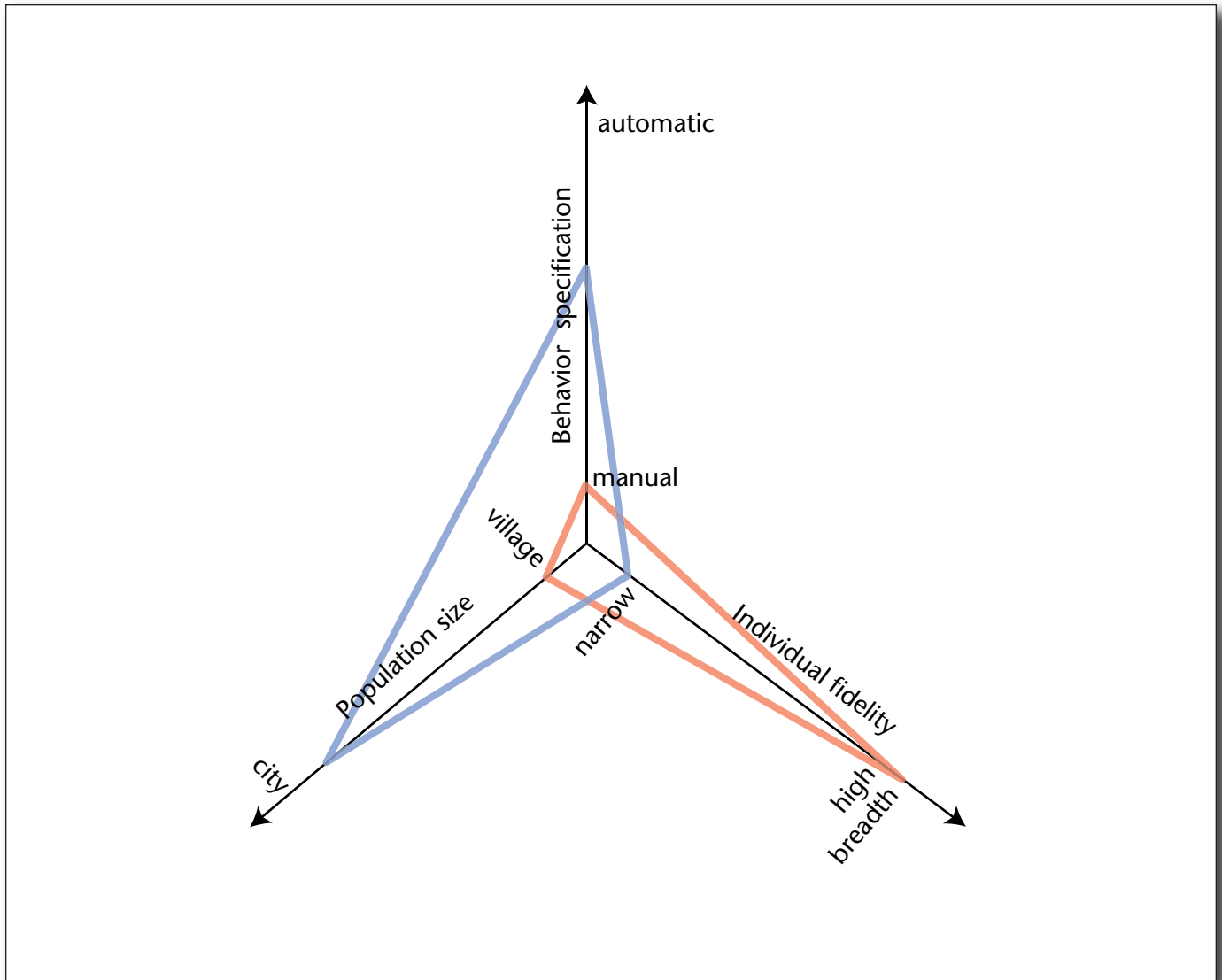
Figure 1. Current Models of Patterns of Life Do Not Scale in All Three Model Dimensions.

models (which may produce unrealistic POL). Researchers must determine which aspects can be automatically generated and how to ease human control over details that cannot be automated.

Authoring POL content can draw from real-world data such as video sources, authoritative input from residents or anthropologists, or cultural artifacts like media. Current research in these fields will support improvements to POL, but important issues for computational modeling remain. Content generalization raises questions of identifying cultural universals, defining cultural differences, and creating relevant model overlays that efficiently capture the differences. Run-time adaptation of the patterns being displayed requires model representations that can influence many individual behaviors with fewer, more abstract directives. Without this abstraction capability, the patterns produced will be unacceptably

dependent on individual decisions, making it difficult to identify core causal patterns.

Artificial Intelligence Solutions

Current population simulations often have limited application or research impact due to a lack of scalability along at least one of the scalability dimensions (figure 1). A city-sized traffic model for the VOP could be synthesized automatically from observed traffic patterns but would disallow a spontaneous street celebration due to the limited knowledge of individual entities. A high-fidelity model of the interpersonal interactions that occur with a few village elders during a military patrol in the VOP would not generalize to a town-sized population due to knowledge engineering costs. Modeling patterns of life that scale along all three dimensions requires new approaches

and algorithms, built on artificial intelligence foundations. We outline three AI research areas on which future solutions could build.

Pattern Recognition

Instantiating an accurate, specific POL model depends on close observation of actual patterns of life. Encoding these observations through knowledge engineering methods is not practical for scalable POL. Automated methods are needed to identify and extract human behavior patterns in order to make the POL specification scalable, especially in the size dimension. AI offers many examples of transforming captured human behavior into useful patterns. Data mining in conjunction with social network analysis can reveal criminal networks from text documents (Al-Zaidy et al. 2012). In computer vision, researchers can classify social interactions or crowd movements (Mehran, Oyama, and Shah 2009).

New challenges in the fields of data mining and pattern recognition for POL will be to extract and interpret the more subtle variations of everyday life, to recognize gestalt patterns in addition to individual actions or events, and to incorporate commonsense reasoning and narrative understanding into diverse settings.

In order for pattern-recognition approaches to identify and then simulate POL regularities and anomalies, researchers need to learn which features to consider despite all the breadth and subtleties of human interaction. Identifying these features and data sources to support them is a nontrivial task. We suggest it may be fruitful initially to study POL in constrained human domains that are limited in the number of possible observable behaviors and latent goals. Examples of problems with existing data sets might include movement patterns of ships, trucks, airplanes, or people within a specific geographic area or network. Studies targeting such domains could develop new metrics and processes for recognizing high-value inputs, outputs, and patterns for POL in general. This work will support POL if it can give insights into modeling new domains that encompass more numerous and varied observations and hidden processes, such as life in the VOP.

Representational Abstraction

A "missing link" in modeling POL is the lack of general, stable representations of population behavior. However, finding and exploiting computationally precise, useful representations at different levels of abstractions is a hallmark of AI, from the study of the processes of mind and structure of games in AI's infancy to recent advances in market abstractions and agency.

POL researchers have defined representations, such as social networks and crowd models, which reflect AI's representational heritage. These results, while important, are comparable to Simon's (1996)

individual watch parts. Progressing directly from these scattered elements to a useful and scalable POL has proven difficult.

We recommend a shift to identifying higher-level abstractions (subsystem representations and rules to compose them) that encapsulate individual patterns and influences, control their interactions, and thus enable their assembly into the working watch of a scalable POL simulation. Hierarchical representations of crowds, groups, and individuals (for example, Musse and Thalmann 2001) suggest a path toward this goal, but significant extension will be required to include the many factors that influence patterns of life.

Behavior Generation

Computational models of human behavior (for example, Anderson and Lebiere [1998], Laird [2012]) and intelligent agents (for example, Wooldridge [2000]) have also been major thrusts in AI. Generative behavior models are needed in order to reify and explore new findings about POL. Researchers who develop generative POL models will contribute to validation and verification of work in the new field and provide important utility to applied users.

In this area, crowd-integrated cognitive models (Pelechano et al. 2005) and agent-based systems that can plan ahead, reason, and react (Laird 2012) are relevant to generating complex behaviors that can both support and potentially control emergent patterns. Prior work in heterogeneous computational modeling of individuals (Lebiere et al. 2002, Wray et al. 2005) may offer initial points of exploration for heterogeneous behavior generation for POL.

In conclusion, contributions from many fields of AI will be welcomed to meet key challenges and realize new opportunities by creating scalable models for patterns of life.

Acknowledgement

References

Aaby, B. G.; Perumalla, K. S.; and Seal, S. K. 2010. Efficient Simulation of Agent-Based Models on Multi-GPU and Multi-Core Clusters. In *Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques*. Brussels, Belgium: Institute for Computer sciences, Social-Informatics, and Telecommunications Engineering.

Al-Zaidy, R.; Fung, B.; Youssef, A. M.; and Fortin, F. 2012. Mining Criminal Networks from Unstructured Text Documents. *Digital Investigation* 8(3–4): 147–160. dx.doi.org/10.1016/j.diin.2011.12.001

AIIDE-14 Is Moving to Raleigh, North Carolina!

The Tenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment
North Carolina State University
Raleigh, North Carolina
October 3–7, 2014

AIIDE-14 is the definitive point of interaction between entertainment software developers interested in AI and academic and industrial AI researchers. The conference is targeted at both the research and commercial communities, promoting AI research and practice in the context of interactive digital entertainment systems with an emphasis on commercial computer and video games. AIIDE-14 will include invited speakers, research and practitioner presentations, playable experiences, project demonstrations, interactive poster sessions, and a doctoral consortium. The program also includes a workshop program, which will be held on the first two days of the conference, October 3–4.

www.aiide.org

Anderson, J. R., and Lebiere, C. 1998. *The Atomic Components of Thought.* Mahwah, NJ: Erlbaum.

Hughes, R. L. 2002. A Continuum Theory for the Flow of Pedestrians. *Transportation Research* Part B 36(6): 507–535. dx.doi.org/10.1016/S0191-2615(01)00015-7

Jones, R. M.; Folsom-Kovarik, J. T.; McLaughlin, P.; and Frederiksen, R. 2014. High-Density Pattern-of-Life Modeling. Unpublished Manuscript.

Laird, J. E. 2012. *The Soar Cognitive Architecture.* Cambridge, MA: MIT Press.

Lebiere, C., Biefeld, E., Archer, R., Archer, S., Allender, L., and Kelley, T. 2002. IMPRINT/ACT-R: Integration of Task Network Modeling Architecture with a Cognitive Architecture. In *Proceedings of the Advanced Technologies Simulation Conference. Military, Government, and Aerospace Simulation Series* 34(3): 13–18. San Diego: Society for Modeling and Simulation International.

Mehran, R.; Oyama, A.; and Shah, M. 2009. Abnormal Crowd Behavior Detection Using Social Force Model. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 935–942. Piscataway, N.J.: Institute for Electrical and Electronics Engineers.

Musse, S. R., and Thalmann, D. 2001. Hierarchical Model for Real Time Simulation of Virtual Human Crowds. *IEEE Transactions on Visualization and Computer Graphics* 7(2): 152–164. dx.doi.org/10.1109/2945.928167

Pelechano, N.; O'Brien, K.; Silverman, B.; and Badler, N. 2005. Crowd Simulation Incorporating Agent Psychological Models, Roles and Communication. Paper presented at the First International Workshop on Crowd Simulation (V-CROWDS '05), Lausanne, Switzerland, 24–25 November.

Preston, S. W. 2012. CIA and the Rule of Law. Harvard Law School, Cambridge, MA. April 10, 2012. Public Address. *Journal of National Security Law and Policy* 6(1).

Schatz, S.; Folsom-Kovarik, J. T.; Bartlett, K.; Wray, R. E.; and Solina, D. 2012. Archetypal Patterns of Life for Military Training Simulations. In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference.* Arlington, VA: National Training and Simulation Association.

Simon, H. A. 1996. *The Sciences of the Artificial.* Cambridge, MA: MIT Press.

Snow, J. 1855. *On the Mode of Communication of Cholera.* London, England: John Churchill. PMCid: PMC2438792

Wooldridge, M. 2000. *Reasoning about Rational Agents.* Cambridge, MA: The MIT Press. PMCid:PMC2244060

Wray, R. E.; van Lent, M.; Beard, J.; and Brobst, P. 2005. The Design Space of Control Options for AIs in Computer Games. Paper presented at the 2005 IJCAI Workshop on Reasoning, Representation, and Learning in Computer Games, Edinburgh, UK, 31 July.

J. T. Folsom-Kovarik is a research scientist in the intelligent training area of Soar Technology, Inc. His research addresses innovative ways adaptive systems can learn about, model, and influence human behavior, cognition, and learning.

Sae Schatz is the chief scientist of MESH Solutions, LLC (a wholly owned subsidiary of DSCI). Her work focuses on education and training research, including instructional technologies, adaptive systems, and modeling and simulation.

Randolph M. Jones is a cofounder and senior artificial intelligence engineer at Soar Technology. He has been developing individual-level and aggregate-level cognitive models for more than 25 years, including several applied models for defense systems.

Kathleen Bartlett is an instructional systems design specialist at DSCI's MESH Solutions Division. She is an award-winning instructor with research experience in a variety of military and civilian settings.

Robert E. Wray is a senior scientist at Soar Technology. His research encompasses many areas of artificial intelligence including agent-based systems and agent architectures, AI in education, machine learning, cognitive science, and knowledge representation and ontology.