Computational exploration of simulations in political science

R. Jordan Crouser, Jeremy G. Freeman, and Remco Chang

A novel method of constructing graphs of agent-based simulation data enables social scientists to better understand complex systems.

Agent-based simulation (ABS) has become an important tool for studying complicated group behaviors in social science.[1] In ABS, large numbers of autonomous entities, or 'agents,' interact with one another and, over time, they influence and are influenced by the agents around them. Given that these simulations are stochastic, that is, they use small random perturbations, they generally have to be run hundreds of times to generate a distribution of sample behavioral patterns. Increased computing power enables scientists to simulate and study increasingly complex systems. For example, ABS is being used to model cooperative behavior,[2] ethnic mobilization and conflict,[3] violence and genocide,[4] and population growth and collapse.[5]

Since these systems are being used to simulate elaborate patterns of human behavior, they require thousands of agents each with a large number of variables to direct its actions and interactions with those around it. Such complexity means that these systems often generate gigabytes of raw data for each simulation. Unfortunately, these very large data sets can prove incredibly costly to interpret and analyze. Lacking appropriate tools and support, scientists studying these systems are driven to oversimplify their models or perform purely numerical analyses that, by design, overlook many of the subtle yet important forces driving behaviors.

To get around this issue, we reframed this process as a graph exploration problem. In other words, we began by asking questions about which configurations in the simulation can be reached from a given starting position. In this context, a graph is a representation of objects (called vertices) and the relationships between those objects. Vertices are unique configurations of the variables controlling the simulation, and the relationship between two configurations is a transition in time. This reframing enables us to employ concepts from graph theory to describe behavioral patterns in these data sets.
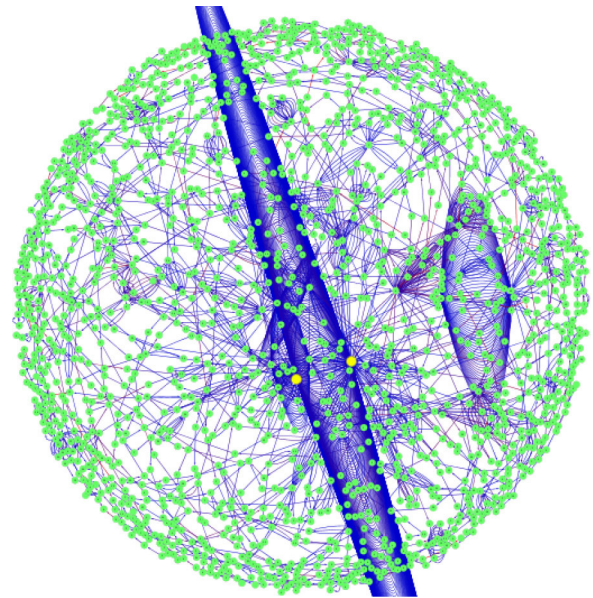

Figure 1. *Visualization of an aggregate temporal graph generated from 100 runs of an agent-based simulation of political hierarchies. The two yellow vertices represent an interesting feature: a highly stable vertex pair (note the high degree of revisitation).*

To begin our analysis, we first constructed a graph formulation called an aggregate temporal graph (ATG) from the data. Here, a vertex represents a unique state in the simulation, and a directed edge represents a transition between two states in one time step. To build the graph, we combined a large set of sample sequences from the agent-based simulation using the following rule: if two sequences have identical values for all variables at some point in their execution, then that state is represented as a single vertex that is shared by both sample sequences. Each of these sequences can be thought of as a pathway through the simulation space (the set of reachable configurations inside the high-dimensional parameter space). The ATG structure differs from traditional graph structures in that it highlights the relationships and transitions between simulation states, rather than

focusing solely on the interactions between individual entities across multiple runs. A simulation state is a configuration of all variables controlling the simulation at a discrete moment in time.

Following the construction of an ATG, we can then begin to look for interesting features in the data sets. These include states or patterns that appear repeatedly in many simulations, future states that are possible from a selected state, and uncommon configurations that may represent unlikely yet significant events in the simulation.

Colleagues at the University of Pennsylvania working in political science used ABS to model political violence and power systems in present-day Thailand using a model called dynamic political hierarchy (DPH).[6] Their simulation modeled the interaction of 31 political identities, tracking their popularity, level of violence, protest, power, and influence. This problem is a perfect example of a case where our graph formulation can be applied to understand simulation results.

We formulated raw data from 1000 runs of their simulation in an ATG containing 14,887 unique vertices and 59,000 edges. Each vertex represents a unique DPH configuration of the 31 interacting political identities, and a directed edge represents a temporal transition between two DPH configurations over the course of one time step. We built a subgraph from the first 100 runs of the simulation (see Figure 1). We were then able to explore the structure and connectivity of different parts of the graph to identify some of its interesting characteristics.

We wanted to determine the set of possible outcomes of a set of starting configurations, and the likelihood of each outcome. Further, we were especially interested to detect any events that were inevitable, given a particular starting point. For example, imagine we begin with X and Y vying for power. From our graph, we see that while X may rise to power, Y is more likely to win out. Either way, we might find it is inevitable that we will see a large amount of protest from Z, who disagrees ideologically with both X and Y and will be unhappy regardless of the outcome.

It is also important to be able to find critical decision points preceding undesirable states. The terminology often used to describe these decision points is 'left of boom,' indicating that appropriate actions could still prevent catastrophe. Our method can help in this regard. Using the ATG built from the DPH simulation data and both well-studied[7] and our own novel graph exploration techniques, we were able to answer these questions in real time. For example, we want to identify states with several possible outcomes, some with a high level of violence and the others relatively stable, and determine what factors influence their outcomes, in order to prevent violent outbreaks in the real world.

By constructing an ATG from simulation data, we were able to easily detect possible, likely, and inevitable events in a well-described simulation space. This information could be used by analysts to better understand the consequences of actions. In our future work, we plan to apply these ideas to similar problems in other domains. Furthermore, we intend to develop robust, domain-specific software tools to support analysts in more fully exploring complex systems, enabling them to better inform decision-makers and effect real-world change.

Author Information

R. Jordan Crouser, Jeremy G. Freeman and Remco Chang
Department of Computer Science
Tufts University
Medford, MA

R. Jordan Crouser is a doctoral candidate, whose research interests include human-computer interaction and visual analytics. His current work focuses on using graph techniques to develop better tools for understanding large-scale political science simulations.

References

1. I. Lustick and D. Miodownik, *Abstractions, ensembles, and virtualizations simplicity and complexity in agent-based modeling*, **Comparative Politics 41** (2), pp. 223–244, 2009.
2. R. Axelrod, **The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration**, Princeton University Press, 1997.
3. A. Srbljinovic, D. Penzar, P. Rodik, and K. Kardov, *An agent-based model of ethnic mobilisation*, **J. Artif. Soc. Soc. Simulat. 6** (1), p. 1, 2003.
4. R. Bhavnani and D. Backer, *Localized ethnic conflict and genocide*, **J. Conflict Resol. 44** (3), pp. 283–306, 2000. doi:10.1177/0022002700044003001
5. R. Axtell, J. Epstein, J. Dean, G. Gumerman, A. Swedlund, J. Harburger, S. Chakravarty, *et al.*, *Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley*, **Proc. Nat'l Acad. Sci. U.S.A. 99** (Suppl. 3), pp. 7275–7279, 2002. doi:10.1073/pnas.092080799
6. I. Lustick, B. Alcorn, M. Garces, and A. Ruvinsky, *From theory to simulation: the dynamic political hierarchy in country virtualization models*, **Am. Pol. Sci. Assoc.**, 2010. http://ssrn.com/abstract=1642003
7. J. L. Gross and J. Yellen, **Graph Theory and its Applications**, ch. 12, CRC Press, 2006.