SAND2019-4189C

Complexity metrics for Agent Based Models of Social Systems *

Kiran Lakkaraju, Asmeret Naugle, Stephen J. Verzi, Laura P. Swiler, Michael Livesay, Christina Warrender, Michael Bernard, and Vicente Romero

Sandia National Laboratories

Abstract. We develop a multi-tiered approach to measure the complexity of agent based models of social systems, incorporating four interacting but complementary aspects of complexity: system intricacy, information theoretic complexity, behavioral capacity and social organization. We apply these metrics on the classic Schelling model of segregation as an example.

1 Introduction

Agent based modeling (ABM) has a long and rich history in studying social phenomena. The benefit of ABM is in developing simulations in which complex patterns emerge. The extent to which underlying micro-processes and the resulting patterns of behavior are similar to the real world is the question of validity and is an extensively studied question [7,2].

We focus on an aligned question, what does it mean for an agent based model to be "complex"? While there has been study of the computational complexity of multi-agent systems (MAS) [19] and the complexity of MAS software [12] there is very little work that studies the complexity of multi-agent systems in a way similar to that of the real world. Thousands of models with widely differing micro processes have explored historical, fictional and futuristic domains. How do we distinguish between different agent based models? How do we quantitatively compare across different models over different domains?

To address these questions, we propose a quantitative, multi-tier definition of complexity that can be used in studying agent based models of social systems. Our metric is founded on insights from complexity theory, the social sciences, and software engineering. By integrating multiple domains in the development of our metric we are able to better capture the multi-faceted nature of complexity.

Using our metric, agent based models of social systems can be quantitatively compared with each other, allowing us to better understand their utility.

* Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA0003525.

2 Ground truth

All agents have an underlying decision process that integrates perceptions of the environment, signals from other agents, and their own goals to determine behavior autonomously. We call this underlying decision model the "ground truth" of the model, since it dictates the behavior of agents in the simulation (the "micro level") and, thusly, the behavior of the simulation as a whole (the "macro level" effects [23]).

As a simple example, consider a 2-D cellular automata. Each cell perceives its neighbor's states and autonomously changes its state in response based on the rules of the model. The decision rule for each cell is the ground truth of the model. If the decision rule changes, the macro level simulation behavior can dramatically change [3].

A variety of models of agent decision making exist such as the Belief-Desire-Intention (BDI) cognitive model [20], or the Partially Observable Markov Decision Process (POMDP) model [18].

Our goal is to identify a simple representation of agent decision making that will capture the following:

- How agents in the simulation make decisions.
- How agents in the simulation interact with each other.
- How agents in the simulation interact with their environments.
- Any environmental factors that influence each other within the simulation.

We will constrain the ground truth to focus only on causal connections between variables and parameters within the model. Causality is defined by interrogation of the decision rules and equational relationships for the model. For example, if variable A is used in the equation/algorithm for calculating variable B, then we say that A causally impacts B.

We represent the ground truth as a graph, Nodes represent variables and functions/aggregations of variables, and edges between nodes represent causal relationships between nodes. For example, if variable A is used in the equation/algorithm for calculating variable B, then we include a link from Node A to Node B.

We allow nodes to represent functions/aggregations over agents or the environment (for instance, a node could be the average value of a state across all agents in the simulation).

In identifying the ground truth, we follow these principles:

- Nodes should be combined where possible. If there are multiple simulation variables that represent similar concepts and have the same causal structure (ie: the same causal influences), then those variables can be represented as a single node.
- Relationships between entities should be represented as simply as possible. For example, the ground truth does not need to represent the entire influence network between agents in the simulation; instead, links can represent types of causal relationships between generic agents.

– The exact functional form of equations and parameterization are not represented in the ground truth. The ground truth diagram is only meant to specify causal relationships.

Section 4 outlines the ground truth for the classic Schelling Segregation model.

3 Proposed multi-tier complexity metrics.

Table 1. Complexity metrics organization

	Not tied to social/behavioral science	Inspired by social/behavioral science
Requires knowledge of system structure	System Intricacy	Behavioral Capacity
Does not require knowledge of system structure	Information-Theoretic	Social Organization

Delineating between simple, intricate (sometimes also referred to as complicated), complex, and chaotic systems is a difficult task. Many definitions of complexity have been proposed in the literature [9,14] but no definition is widely accepted. We are focused on assessing the complexity of models of social systems. To measure the complexity of social simulations, we have identified a multi-tiered suite of metrics that captures different elements of complexity. Using a carefully chosen combination of methods, we can gain a deeper and more nuanced understanding of simulation complexity than could be achieved with a single metric.

The complexity metrics are organized along the two dimensions in Table 1. The first dimension (rows) differentiates between metrics that require knowledge of the system structure (i.e., ground truth) of a simulation and those that do not. Metrics that require knowledge of the system structure may be useful for causal simulations, but we generally do not have knowledge of the causal structure of real-world systems. However, if we can develop methods to infer this causal structure, these complexity metrics may apply on real-world systems.

The second dimension (columns) relates to the original intended application space of the metric. The right-hand column includes metrics that are inspired by the social and behavioral sciences, while metrics in the left-hand column measure more abstract properties of the simulation, and might be inspired by other application spaces or might be purely mathematical. We focus on the social and behavioral sciences since our focus is on modeling of social systems.

The four metrics are described in more detail below.

3.1 System Intricacy

Measures of system intricacy capture the complexity of a simulation's causal structure, or ground truth. These metrics are inspired by the notion that the more components and causal relationships a system has, the more complicated it is.

System intricacy is intimately tied to the causal structures, processes and interactions that determine the dynamics of the system. One approach for measuring system intricacy in simulations is to evaluate the complexity of the structure of the underlying software implementation, however these are not pure metrics of the system. We evaluate the system intricacy of a simulation by evaluating the simulation's ground truth.

Cyclomatic complexity (initially proposed in [13]) was initially developed for studying the complexity of software, however it has since been used in other domains [16]. We use it as a concise summary of the complicatedness of a simulations ground truth.

Cyclomatic complexity (M), captures the interconnectedness of a graph by counting the nodes (N), edges (E) and the number of connected components (P) in a graph:

$$M = E - N + 2P$$

3.2 Behavioral Capacity

Behavioral capacity measures capture the potential for rich and diverse interaction potential among agents in a system. The underlying hypothesis is that complex simulation of social processes will include significant and varied interaction between agents. Humans participate in a wide variety of groups, at multiple scales (from country membership to family groups). A complexity measure that captures this will capture an important part of human behavior.

A variety of metrics can be used to represent behavioral capacity of a social simulation, such as the number of interactions between agents, or the number of groups an agent participates in.

We focus on a measure that explicitly counts the number of relationships an agent has, the number of differentiated relationships, because of its intuitive appeal, ability to quantify, and prior work in the literature [1].

Intuitively, an agent that has multiple different types of relationships must juggle different goals and needs. An individual must do the same when they interact with a shopkeeper vs. family member. The difference in relationship can naturally track that of group membership.

Quantification of this measure can be done by viewing the ground truth of the model. Since the ground truth specifies all interactions between agents we should see evidence of differentiated relationships as types of influences and interactions agents can have with each other.

3.3 Information-Theoretic

Information-theoretic complexity measures capture information content related to the dynamics of a system. These metrics are inspired by the notion that a more complex system will generate more information over time. These metrics account for uncertainty, and are calculated using a systems (or simulations) input and/or output data (see [22] for a review). Information-theoretic complexity metrics have been developed and used in several fields. These metrics may not always capture our intuition of complexity; for example, these measures might consider randomness to be a form of complexity, since uncertainty and information content are entangled. We address this by considering information-theoretic complexity metrics in conjunction with the other three metric categories. These metrics are calculated using data directly from the social system or simulation results.

Many information theoretic metrics have been proposed in the literature, such as entropy [4], mutual information [4], autocorrelation [11], and compression ratios [10]. We focus on forecasting complexity (C) [22], which captures the minimum amount of information (H) needed for optimal prediction within a time-series, where part of the time-series, X^-, is used to predict the rest, X^+(such that $X = (X^-, X^+)$), using a model f in M (where M is a specific space of models):

$$C = \min_{f \in M} H(f(X^-))$$

Forecasting complexity captures an intuitive notion of complexity based on prediction, but it is hard to compute and requires a space of models (M) to search. Here, we use an approximation to forecast complexity involving compression ratio which itself is an approximation to normalized information distance [10] (which we call approximate NID).

For the information theoretic complexity on a time series of data, normalized information distance between the past and future information is defined as follows.

$$NID(x, y) = \frac{\max\{K(x|y), K(y|x)\}}{\max\{K(x), K(y)\}},$$

We approximate this using split points computed over the entire time series giving a series of approximate NID complexities. This series of complexities is then averaged. The equation is given by

$$\text{approximate}NID(x, y) = \frac{\max\{Z(x|y), Z(y|x)\}}{\max\{Z(x), Z(y)\}},$$

where Z takes time-series information and gives the size of information after conditional compression. The LempelZiv-Markov chain algorithm (LZMA) compression is used because it takes and creates dictionaries for compression as it compresses and uses these dictionaries to compress new strings processed in the future, thus implementing a notion of conditional compression used here in approximate NID.

Let $S = \{S_1, ..., S_n\}$ be the data series. Then we are interested in the series approximate$NID(S_t, S_{t+1})$ for $t = 1, 2, ..., n-1$.

In theory approximate$NID(x, y) \in [0, 1]$, because in theory $Z(x|y) < Z(x)$ and Z is positive valued. The normalized p-norm is $(\sum_{i=1}^{n} f(i)^p \frac{1}{n})^{1/p}$. If $p = 1$ then indeed the mean is given. If $p = \infty$ (that is the limit as p gets arbitraily large), then the maximum is given. Note that $\lim_{p\to\infty} \frac{1}{n}^{1/p} = 1$. Since the information theoretic complexity is the normalized 2-norm it is in $[0, 1]$ and it is moderately influenced by the maximum and the average. But this complexity is resilient to maximums which are outliers, as well as large sequences of constant values which may be due to poor choice of cut-off times for the simulation.

3.4 Social Organization

Measures of social organization capture information about how individuals form groups, how groups combine to form larger groups, and how individuals and groups interact. These metrics are inspired by the idea that complex social systems demonstrate emergent hierarchical organization and complicated interactions between individuals and groups [17]. This category of metrics addresses the interaction between different levels of analysis, i.e., micro, meso and macro scale patterns within the system [6]. These metrics will be calculated using simulation results – the characteristics, states, and actions of agents during the simulation.

To preserve generality, we focus on measures that apply to a social network generated by a simulation. The social network represents interactions between agents in a simulation. A node in the social network represents an agent, and an edge represents interaction. We can extract interactions between agents from simulation output to create a social network.

Quantitative characteristics of a social network have been used to characterize real world social systems ([5]). Existing literature suggests a variety of means to capture different aspects of social networks, including clustering coefficients, community detection algorithms, and centrality measures.

We focus on measures that can capture the hierarchy within a social system. Hierarchies are an important concept in the social and physical sciences, and have been considered a fundamental characteristic of complex systems [17]. To quantify the hierarchy within a simulation we focus on the Global Reaching Centrality measure (GRC) as defined in [15], which uses the local reach centrality.

Let $C_R(i)$ be the local reach centrality of node I, the proportion of nodes that can be reached from node I via outgoing edges. Then the Global Reach Centrality (GRC) is defined as:

$$GRC = \sum_{i \in V} \frac{[C_R^{\max} - C_R(i)]}{(N-1)}$$

This definition can be easily extended to undirected networks by considering the weights on the edges (with a default weight of 1) for every edge. The GRC can range from 0 to 1, with a higher value indicating a higher level of hierarchy in the social network [15].

4 Example application: Schelling segregation model

As an illustrative example, and to highlight potential difficulties, we consider applying our complexity metrics to the classic Schelling segregation model [21]. Our goal is not to extensively evaluate the Schelling model, but rather to highlight the promise, and understand the pitfalls, of our multi-tier complexity metric.

We use the NetLogo implementation of the Schelling model [24]. Agents are characterized by a color that is fixed throughout the simulation and are located on a 2-d lattice. Only one agent can be at any single point on the lattice.

Agents have a preference to be with like-minded (i.e., same color) agents. The premise is that if an individual has a preference to be in a neighborhood where a larger percent of neighbors have similar traits to themselves than their current neighborhood, then and only then are they motivated to move to a new residence (empty lattice site).

By leaving, the agent has positively reinforced the current dominating trait; while on the other hand, this agent's presence at their new location reinforces their own trait at the new location. Such an act reinforces the average mindset of the neighborhood, causing any unlike-minded neighbors to be even more outnumbered, hence they have reinforced segregation on multiple fronts. Due to this positive feedback, even a slight intolerance (such as the need for 26% like-minded neighbors) could potentially lead to highly segregated regions.

We consider a simple version of the Schelling model here, but there has been extensive study of this model, see [8]

Ground truth for the Schelling model Agents in the Schelling model have the following characteristics (fixed features of the agent) and states (dynamic features of an agent):

Characteristics :
- **Color** An actor has a color that is fixed throughout the simulation.
- **Preference Ratio** The percentage of neighbors of an actor that should share the same color, denoted as p_{pref}.

Behaviors :
- **Change Location** An actor can take the action to change it's location to another empty location in the grid.

States :
- **Location** Location of an actor on the grid.

Let p_{matching} be the fraction of neighbors of an actor that have the same color as the actor. On a time step, the decision rule an actor executes is the following:

$$\textbf{if } p_{\text{matching}} < p_{\text{pref}} \textbf{ then } \text{Actor Changes Location}$$

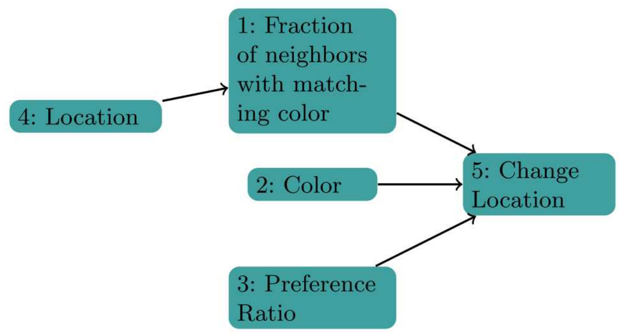

Fig. 1. Ground Truth for the Schelling model

Fig. 1 is the ground truth for the Schelling segregation model.

Node 1 is an aggregation of the characteristics of neighboring agents. Node 2 is the color of an agent. Node 3 is the preference ratio for an agent. Node 4 is the location of an agent, and Node 5 is the behavior an agent undertakes.

In this specific implementation, all agents have the same preference ratio. Even if agents had different preference ratios, the ground truth would not change as it only captures the fact that there is a relationship between preference ratio and the behavior to change a location.

5 Parameters and Types of Behaviors

The Information-Theoretic and Social Organization measures are computed on simulation output and are thus dependent upon the parameterization of the simulation. There are several parameters in the Schelling model that can influence the system behavior.

Density The fraction of locations on the 2-D lattice on which an agent resides. If the density of the population is too low then all the agents may have no desire to move, as they have no neighbors. If the density were 100% then the agents could not move. Moreover, given a good value of density there exist high values of preference so that the agents never settle; likewise there are such low values of preference that they settle immediately.

Preference Ratio Setting If the preference is above 80% (and even above 75% for most densities) then enough agents choose to move that the simulation becomes chaotic and uninformative. If the preference ratio is too low, agents do not desire to move at any time and the simulation exhibits no dynamics.

The range of density in NetLogo's segregation model is $50\% - 99\%$, all of which are determined to be reasonable. For densities less than 50%, the simulation converges too quickly or gets stuck in a random cycle. When the preference for the agents is too high, then on the boundary of the segregated regions the agents will choose to move and move randomly, making the boundary grow. Eventually, the entire population is moving randomly, rarely staying in any location.

6 Application and discussion of Complexity Metrics

6.1 System Intricacy & Behavioral Capacity

For the Schelling model, the system intricacy is $M = 4 - 5 + 2 * 1 = 1$. The behavioral capacity is simply the number of differentiated relationships, which is determined to be 1.

This aligns with intuition and general perception. In fact, its importance derives from the fact that so few elements are needed to produce what is thought to be a complex pattern of behavior.

All agents in the Schelling model interact with each other in the same way, by evaluating their color. A counter argument would be that since color determines action, and there is a different action for agents that are of a different color, that would indicate a different relationship. However, note that the action of an agent does not have a subject – no agent does anything to another agent. This is the underlying characteristic of a differentiating relationship, one in which there are different actions towards different agents, of which there are none here.

6.2 Information Theoretic Complexity

To apply our information theoretic measure, we need to identify the appropriate information to collect at each time step. In this simple model we can use the states of every location on the grid as a representation of the simulation at each time step.

Table 2 shows the information theoretic complexity values for a variety of parameter settings, chosen to highlight different behaviors.

This metric aligns with our intuition. When there is near instant convergence, the information theoretic complexity is low (0.38 for density = 50% and preference ratio at 50%) vs. situations in which there is lots of movement (density = 99% and preference ratio of 60%).

We acknowledge a weak correlation with the number of timesteps, but note that our 2-norm method alleviates some of that. Compare the values for density=90% vs. density=50% for the preference ratio value of 50%, we can see that even with fewer time steps to converge the information theoretic measure was higher.

Density	Preference 50%	Preference 60%
99%	0.64 (35)	0.96 (1001)
90%	0.59 (19)	0.72 (99)
80%	0.46 (25)	0.63 (33)
70%	0.49 (21)	0.58 (30)
60%	0.47 (20)	0.59 (23)
50%	0.38 (28)	0.62 (19)

Table 2. Information Theoretic complexity examples at convergence (time steps to convergence indicated in parentheses, simulation stopped at 1000 timesteps if not converged), given density and preference.

6.3 Social Organization

We define the social network for the Schelling model in the following way. An unweighted edge is established between two agents if they were neighbors at any point during the simulation. The global reaching centrality (GRC) was calculated on this network, see table 3.

Density	Preference 50%	Preference 60%	Preference 70%
99%	0.1648	0.1628	0.1583
90%	0.0931	0.1012	0.0898
80%	0.0880	0.1129	0.0537
70%	0.0801	0.0979	0.0494
60%	0.0825	0.0954	0.0654
50%	0.0870	0.0910	0.0717

Table 3. Global Reaching Centrality (GRC) examples after 500 time steps, given density and preference. Given a particular density, it would appear that the GRC is at its highest for the preferences which converge the most quickly. This table does not include parameters for when the agents move randomly forever.

The GRC measure of the simulation runs can be compared to real world examples. [15] calculates the GRC for a variety of real world graphs. Food webs have a high GRC. Surprisingly, trust in an organization has low GRC scores. Our results show quite a low value of GRC for the parameter setting, indicating that the social network we defined based on neighbors is not very hierarchical. This makes sense, as agents move around in the grid. We can also notice a pattern of decreased hierarchy as the density increases. There should be a correlation between increase in density and agent moving (if they desire to move, but can't find a place to move, they will continue to desire to move).

System Intricacy	Behavioral Capacity
1	1
Info.-Theoretic	**Social Organization**
0.59	0.10

Table 4. Summary of complexity metric for the Schelling model. Info. Theoreticand Social Organization values are means over parameters settings defined in Table 2 and Table 3, respectively.

7 Discussion & Conclusion

Table 4 summarizes our assessment of the Schelling model using the multi-tier complexity metrics. In any simulation there will be critical parameters that can impact simulation behavior. Two of our metrics are sensitive to those parameters since they are dependent upon the simulation output. As we have done, to appropriately use our complexity metrics it is necessary to characterize behavior in the parameter spaces. This may require significant resources (computational time).

Each individual metric, taken by itself, will have some drawbacks. However, when considered together we believe they capture important classes of models. As a guide, consider classifying systems based on the system structure metrics (System Intricacy and Behavioral Capacity) vs. the metrics that are based on simulation output (Information-Theoretic and Social Organization). We can determine four different cases based on whether the values for system structure and simulation output, respectively, are "low" or "high":

Low-Low These are systems which are not intricate and do not produce unpredictable behavior. A simple linear system, such as $x_t = 1.1 * x_{t-1}$ may be an example of this.

High-Low These are intricate systems that produce simple behavior. This is interesting as the intuition is that an increase in the complexity of the causal structure of a model should result in an increase in complexity. However, these systems do not exhibit this characteristic.

Low-High These are systems, where a simple, non-intricate set of rules can determine complex behavior. Examples abound of these types of models, for instance Rule 110 in the Cellular Automata literature is a simple decision rule but is shown to be Turing complete [3].

High-High These are systems that capture our intuitive notion that more structurally complex systems will have higher complexity. Most real world system, which will contain a multitude of entities interacting over continuous time and space would fall into this category.

We have developed these metrics to be used in conjunction with each other. Together, they can help characterize and organize different model.

References

1. Thore J Bergman and Jacinta C Beehner. Measuring social complexity. *Animal Behaviour*, 103:203–209, May 2015.
2. Carlo Bianchi, Pasquale Cirillo, Mauro Gallegati, and Pietro Vagliasindi. Validating and Calibrating Agent-Based Models: A Case Study. *Computational Economics*, 30(3):245–264, 2007.
3. Keith C. Clarke. Cellular Automata and Agent-Based Models. In Manfred M. Fischer and Peter Nijkamp, editors, *Handbook of Regional Science*, pages 1217–1233. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
4. Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, New York, NY, USA, 2006.
5. R. I. M. Dunbar, Valerio Arnaboldi, Marco Conti, and Andrea Passarella. The structure of online social networks mirrors those in the offline world. *Social Networks*, 43:39–47, October 2015.
6. Jessica C. Flack and David C. Krakauer. Challenges for complexity measures: A perspective from social dynamics and collective social computation. *Chaos (Woodbury, N.Y.)*, 21(3):037108, September 2011.
7. Mattia Guerini and Alessio Moneta. A method for agent-based models validation. *Journal of Economic Dynamics and Control*, 82:125–141, September 2017.
8. Erez Hatna and Itzhak Benenson. The Schelling Model of Ethnic Residential Dynamics: Beyond the Integrated - Segregated Dichotomy of Patterns. *Journal of Artificial Societies and Social Simulation*, 15(1):6, 2010.
9. James Ladyman, James Lambert, and Karoline Wiesner. What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67, Jan 2013.
10. Ming Li, Xin Chen, Xin Li, Bin Ma, and P. M. B. Vitanyi. The similarity metric. *IEEE Transactions on Information Theory*, 50(12):3250–3264, Dec 2004.
11. Benoit B. Mandelbrot. *Fractals and Scaling In Finance: Discontinuity, Concentration, Risk*. Springer Publishing Company, Incorporated, 1st edition, 2010.
12. Toufik Marir, Farid Mokhati, Hassina Bouchelaghem-Seridi, and Zouheyr Tamrabet. Complexity Measurement of Multi-Agent Systems. In *Multiagent System Technologies*, pages 188–201. Springer, Cham, Cham, September 2014.
13. T.J. McCabe. A Complexity Measure. *IEEE Transactions on Software Engineering*, SE-2(4):308–320, December 1976.
14. Melanie Mitchell and Mark Newman. Complex Systems Theory and Evolution. In *Encyclopedia of Evolution*. Oxford University Press, 2002.
15. Enys Mones, Lilla Vicsek, and Tamás Vicsek. Hierarchy Measure for Complex Networks. *PLOS ONE*, 7(3):e33799, March 2012.
16. Fivos Papadimitriou. Artificial intelligence in modelling the complexity of mediterranean landscape transformations. *Comput. Electron. Agric.*, 81:87–96, February 2012.
17. Denise Pumain. *Hierarchy in Natural and Social Sciences*. Springer Science & Business Media, November 2005. Google-Books-ID: lq3jgwe96H8C.
18. David V Pynadath and Stacy Marsella. PsychSim - Modeling Theory of Mind with Decision-Theoretic Agents. In *International Joint Conference on Artificial Intelligence*, 2005.
19. Zinovi Rabinovich, Claudia V Goldman, and Jeffrey S Rosenschein. *The complexity of multiagent systems: the price of silence*. the price of silence. ACM, New York, New York, USA, July 2003.

20. Anand S. Rao and Michael P. Georgeff. BDI Agents: From Theory to Practice. In *In Proceedings of the First International Conference on Multi-Agent Systems (icmas-95*, pages 312–319, 1995.
21. Thomas C. Schelling. Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186, July 1971.
22. Cosma Rohilla Shalizi. Methods and Techniques of Complex Systems Science: An Overview. In Thomas S Deisboeck and J Yasha Kresh, editors, *Complex Systems Science in Biomedicine*, pages 33–114. Springer US, Boston, MA, 2006.
23. Edward Bishop Smith and William Rand. Simulating Macro-Level Effects from Micro-Level Observations. *Management Science*, 64(11):5405–5421, December 2017.
24. Uri Wilensky. Netlogo 4.0. 4. 2008.