

Understanding Emergent Social Phenomena

Methods, Tools, and Applications for Agent-Based Modeling

László Gulyás

Theses of the Doctoral Dissertation

Supervisors:

Prof. George Kampis

Department of History and

Philosophy of Science

Loránd Eötvös University

Budapest, Hungary

József Váncza, Ph.D.

Computer and Automation

Research Institute

Hungarian Academy of Sciences

Budapest, Hungary

Ph.D. Program in Informatics,

Ph.D. School of Informatics,

Loránd Eötvös University, Budapest

Program Director: Prof. János Demetrovics

Budapest, 2005.

Introduction

The application of computational modeling and simulation is gaining momentum in the social sciences, including economics, sociology, and political science. In these fields, it is often viewed as the ‘third symbol system’, an alternative to or ‘middle ground’ between linguistic (verbal) and mathematical modeling. Computational models retain much of the flexibility of linguistic models, yet, they enforce precision and consistency. The three ‘symbol systems’ complement each other: the three methods are often applied to the same problem, thus strengthening our scientific understanding. Nonetheless, some problems yield more easily to one approach than to the other.

Agent-based modeling is a new branch of computer simulation, especially suited for the modeling of complex social systems. Its main tenet is to model the individual, together with its imperfections (e.g., limited cognitive or computational abilities), its idiosyncrasies, and personal interactions. Thus, the approach builds the model from ‘the bottom-up’, focusing mostly on micro rules and seeking the understanding of the emergence of macro behavior.

Agent-based models are, as is the case with any scientific model, drastic simplifications of reality. In particular, as the agents of such models often represent human beings, it is easy to feel objections towards such ‘caricatures’ of humanity. This problem is almost entirely missing from the natural and engineering sciences. Therefore, special care must be taken in preparing these models and in interpreting the results gained. This is why agent-based models normally aim at deepening our understanding of a given system, and providing existence proofs and sufficient conditions for certain phenomena, even though, the method is perfectly capable of developing quantitative predictions. This standpoint is not as uncommon in the social sciences as it is in the natural and engineering sciences. Nonetheless, the understanding gained from agent-based models can be incorporated in research using a different methodology, for example, in the development of a predictive model. Therefore, agent-based models are also viewed by some as a ‘stepping stone’ for predictive models.

The current trend of applying agent-based models in the social sciences started in the mid-1990s. Being a computational method for researchers in the social sciences who are normally not well-skilled in programming, agent-based modeling also offers significant challenges for practitioners of computer science and information technology. In 1996, when the work summarized in this thesis has started, agent-based modeling and simulation had no standards, no widely adopted methodology, and no well-articulated best-practices. Only a single dedicated toolset existed (Swarm), and even that only in a prototype version. Most agent-based models were written using ad hoc programming techniques and languages. Due to the lack of implementation and publishing standards, these models were ‘black boxes’, their results were impossible to check or replicate. The last decade has seen an intensive, yet unfinished development of the method, both in understanding the advantages and limitations of the approach and in furthering the (both technical and methodological) toolset of agent-based modeling. The current practice consists of following published guidelines in designing the models and selecting among a set of well-established simulation packages. Applying formalized best practices in implementing the simulation is also becoming common, which helps in publishing a model description that fully specifies the behavior, yet many times more concise than the source code.

In this dissertation I summarize my contributions to the development and application of the agent-based modeling method. I introduce a framework that helps classifying modeling approaches; I present a programming language for agent-based modeling; and I formalize a set of implementation ‘best practices’. I also present two applied models using the above results: one in spatial economics and one in econometrics. The former gives an explanation for an empirical regularity found in statistical surveys of US metropolitan areas, while the latter shows the importance of social networks in aggregate choices of groups.

Outline of the dissertation

The dissertation is structured in three parts, each of them containing two chapters. In the first part I give an introduction to agent-based modeling and simulation. In the opening chapter I provide an overview of the method, addressing questions related to the methodological-philosophical position of the approach. This is followed by *Chapter 2*, where I define a framework that illustrates the implementation approach of

agent-based modeling. Specifically, the framework yields four *implementation strategies* that highlight differences between so-called ‘equation-based’ and ‘agent-based’ modeling. I argue that the agent-based character of a simulation is rather a *matter of degree* than a binary property.

In the second part I summarize my contributions to improve the toolset of agent-based modeling. I present two groups of computational tools to assist modelers in developing agent-based simulations. In *Chapter 3* I present the Multi-Agent Modeling Language (MAML), the first programming language dedicated to the development of agent-based simulations. Applying *software metrics* I compare MAML models to their functionally equivalent pairs written in Swarm. I show that, on average, MAML programs are 32% less complex and their development requires 44% less *effort*. In *Chapter 4* I propose *design patterns* to describe implementation practices common in agent-based simulations, and I introduce three patterns (the Simple Model, the Model-Agent-Environment, and the Relational Interaction patterns) that help developing more robust and methodologically correct simulations.

The third part is dedicated to applications of the method and of the tools discussed in the first two parts of the thesis. In *Chapter 5* I present an agent-based model of endogenous city formation that is capable of replicating the rank-size regularity found in US survey data, given sufficient heterogeneity is present in the agents’ information access. In *Chapter 6* I introduce another model, one that studies social influence in a setting of repeated discrete choices. Focusing on the structure of the underlying social influence network, I conclude that on small-world networks the system level outcome can be approximated by the, theoretically simpler and easier to handle, fully connected case.

New results

The contributions of this work to the advance of agent-based modeling and simulation have three aspects, which follow the structure of the dissertation itself: theoretical and conceptual refinement of the method, computational tools and methods, and applications of the methodology.

Thesis I

I introduced a conceptual framework that helps classifying computational simulations. It yields four implementation strategies ranging from the ‘equation-based’ to ‘agent-based’ modeling.

Agent-based simulation is often contrasted to traditional, ‘equation-based’ modeling and simulation. While the latter approach is not properly termed as all computational models are, in fact, ‘equation-based’; significant differences exist between these approaches. The proposed conceptual framework highlights these differences and shows that more than these two choices are possible.

The framework consists of five conceptual *levels*: the level of consideration (LoC), the level of representation (LoR), the level of dynamics (LoD), the level of information (LoF), and the level of interest (LoI). Specifying these levels in case of a particular modeling problem is at the discretion of the modeler, but equating or differentiating these levels results in different *implementation strategies*. Working with a case study I discuss the possible consequences of these strategies, ranging from an ‘equation-based’ approach to an ‘agent-based’ strategy.

Thesis II

I prepared the general definition and directed the development of the Multi-Agent Modeling Language (MAML), the first programming language dedicated to agent-based simulation. Applying a series of *software metrics* I showed that, on average, MAML models are 32% simpler and their development requires 44% less *effort* than functionally equivalent programs in Swarm.

The Multi-Agent Modeling Language is a *domain-specific* programming language, in the sense that it is fully dedicated to agent-based simulation. Yet, unlike many other modeling platforms, MAML is unbiased towards the problem domain the particular agent-based model is aimed at. Moreover, it makes no restrictions on the structure of the model or that of the agents. MAML is equally useful, among others, for spatial, cellular automata like models, for network simulations, or for models of agents with complex cognitive capabilities. These latter properties make it *domain-independent* with respect to the problem domain and the model structure.

MAML was developed for non-professional programmers. Therefore, the main motivations in its design were easing the creation of agent-based simulations and making the modeling decisions explicit and easy to follow. To this end, MAML offers keywords and language constructs specific to computer simulation and agent-based modeling.

MAML models are significantly more concise than alternative contemporary implementations. I studied this observation in detail by the application of a series of *software metrics*. My results show that, on average, MAML models are 32% simpler and their development requires about 44% less *effort* than their functional equivalents in Swarm. With its well-defined modeling constructs MAML has also contributed to the emerging standardization and methodological maturity of the practice of agent-based modeling.

The language's first, reference compiler was developed by Sándor Fazekas under my supervision.

Thesis III

I proposed the application of *design patterns*, a software engineering concept, to describe common practices in implementing agent-based simulations. Building on this, I introduced three patterns (the Simple Model, the Model-Agent-Environment, and the Relational Interaction patterns) that help developing more robust, methodologically correct simulations.

Computational simulations and thus agent-based models are computer programs. Therefore, they involve programming work that modelers, especially students of the social sciences, often do not excel in. This fact contributed significantly to the non-standard, often ad hoc nature of agent-based models at the advent of the method. With time, best practices started to emerge and spread, but in an unordered fashion, mostly in personal communications, following social links.

Similar problems were faced earlier by software engineers in professional software development. In response, the concept of *design patterns* was developed to provide a standardized way to describe and communicate application-independent 'best practices' in design and implementation.

In order to strengthen the methodological soundness of the method I proposed to apply *design patterns* in the field of agent-based modeling and simulation. By nature, when the concept is applied to the specific field of agent-based modeling, it undergoes a slight alteration. First and foremost, application-independence is understood in the context of agent-based simulation. That is, the standardized solution is *always* related to agent-based models, but it is independent of the given model at hand. This also implies that agent-based modeling patterns are often more complex than traditional software engineering patterns, since they deal with the specifics of agent-based simulation. Nonetheless, some otherwise common implementation practices are also elevated to the pattern level in the context of computational models, since they provide novel methodological advice to modelers.

I also introduced three novel patterns: the Simple Model pattern, the Model-Agent-Environment pattern, and the Relational Interaction pattern. The first formalizes a solution for simulations with simple event structures, mainly for modelers who are novices to agent-based modeling. The second helps building simulations that are flexible with respect to different interaction topologies, and can effortlessly be checked against various communication structures. Finally, the Relational Interaction pattern furthers the goals of the Model-Agent-Environment pattern and provides an implementation-independent view of the interaction topology that is able to handle multiple, parallel structures easily.

Thesis IV

I created a model of endogenous city formation where bounded rationality is modeled as limited access to information (i.e., vision). I showed that if sufficient heterogeneity is present in this property of the agents, the rank-size distribution of the cities generated by the model replicates the statistical properties found in US survey data.

The rank-size distribution of US metropolitan areas displays an interesting regularity. According to US statistical surveys, the population of a city is inversely proportional to its statistical rank. This empirical regularity has received special attention in the economics and complexity science literature, yet, arguably no plausible explanation has been proposed so far considering individual-level decisions.

In this application of the agent-based modeling and simulation method, I built on the case study developed for Thesis I. I presented a generative model that restricts agents' access to information. Using this model I successfully replicated the rank-size regularity found in US survey data.

The model involved in this thesis resulted from my collaboration with Yuri Mansury. The involved economic concepts and constructs are his inventions. My contributions amount to the following. I developed a novel agent-based model to replicate the power law rank-size distribution of US cities. I explicitly modeled bounded rationality in the form of limited information access (vision), and showed that if sufficient heterogeneity is present in the agents' vision parameter, the model indeed generates a power law rank-size distribution.

Thesis V

I proposed to study social influence in a class of discrete choice models using well-understood generative models of social networks. I presented an agent-based version of the given model class, with special focus on the agents' interaction topology. I showed that the phase transition in the underlying network model carries through to the aggregate choice made by the population both in the case of Erdős-Rényi and Watts-Strogatz networks.

Econometric discrete choice analysis is a standard tool in many modeling domains, including transportation demand modeling, the practical motivation for this work. It allows for the computation of individual choice probabilities, given the heterogeneous agents' evaluation of alternatives. However, as Dugundji points out, the method is fundamentally grounded in individual choice, while it seems plausible to assume that social networks influence individuals. Therefore, recent works in the field have focused on incorporating the interdependence of various decision-makers' choices, i.e., social influence.

In this application I built on the ability of the agent-based method to explicitly model the interaction among the agents. Using this capability, I presented an agent-based version of a class of discrete choice models, in which the interaction structure is generated by well-understood models of social networks. I studied how these classes of networks effect the system-level, aggregate decision made by the agent population.

I showed that the phase transition in the underlying network model carries through to the aggregate choice made by the population both in the case of Erdős-Rényi and Watts-Strogatz networks. This finding has a practical importance in discrete choice analysis, as empirical studies may test whether real world networks belong to a given network class, while they may never be able to chart the actual social network. In particular, on small-world networks the system level outcome can be approximated by the, theoretically simpler and easier to handle, fully connected case.

The agent-based model involved in this thesis resulted from my collaboration with Elenna R. Dugundji. She has brought both the problem of social influence and the discrete choice models of Aoki, and Brock and Durlauf to my attention. The agent-based version was also built on her analytical results and on her extensions to the named models. My contributions amounted to the following. I proposed the application of well-understood generative models of social networks to represent social influence. I formulated the agent-based version of the model, and it was me who showed that the phase transition in the aggregate choice coincides with the evolution of the giant component in case of Erdős-Rényi networks, and with the collapse of the average path length in case of Watts-Strogatz networks.

Publications on the topic of the dissertation

- [1] Cederman, L.-E. and Gulyás, L.: "Tutorial Sequence for RePast: An Iterated Prisoner's Dilemma Game", Electronic Source, 2001. <http://www.santafe.edu/sfi/education/csss/files02/RepastTutorial.pdf>
- [2] Dugundji, E. R. and Gulyás, L.: "Empirical Estimation and Multi-Agent Based Simulation of a Discrete Choice Model with Network Interaction Effects: An Exploration of the Role of Global versus Local and Social versus Spatial Networks in Transportation Mode Choice Behavior in the Netherlands", In *Proceedings of Computers in Urban Planning and Urban Management Conference (CUPUM 2003)*, 27-29 May, 2003, To Appear, Sendai City, Japan. 2003.
- [3] Dugundji, E. R. and Gulyás, L.: "Socio-Dynamic Discrete Choice on Networks: Multi-agent simulation of social & spatial interaction effects in residential & transportation mode choice", In *Proceedings of the 10th World Conference on Transport Research*, Istanbul, Turkey, 4-8 July 2004.
- [4] Dugundji, E. R. and Gulyás, L.: "Dynamic Discrete Behavioral Choice on Networks: The Nested Logit Model – Analytical Results and Empirical Application", In *Proceedings of XXIV International Sunbelt Social Network Conference*, To Appear, Portorož, Slovenia. 2004.

- [5] Dugundji, E. R. and Gulyás, L.: "Discrete Choices on Networks: The Nested Logit Model – Analytical and Simulation Results", Manuscript, *Santa Fe Institute Working Paper*, Santa Fe, NM, USA. To Appear.
- [6] Dugundji, E. R. and Gulyás, L.: "Socio-Dynamic Discrete Choice in an Empirical Example on Intercity Travel Demand: An Agent-Based Approach, Analytical Benchmark and Some Issues in Estimation", In *Proceedings of Workshop on Modelling Urban Social Dynamics*, Institute of Advanced Studies, University of Surrey, Guildford, UK, To Appear, 2005.
- [7] Gulyás, L.: "On the Transition to Agent-Based Modeling: Implementation Strategies From Variables To Agents", *Social Science Computer Review*, Vol. 20:4, Winter 2002, pp. 389-399, 2002.
- [8] Gulyás, L.: "Relational Agent Models – A Framework for Modular and Topology-Independent Simulations", In *Proceedings of Seventh Annual Swarm Users/Researchers Conference (SwarmFest 2003)*, Notre Dame, Indiana. 2003.
- [9] Gulyás, L.: "Review of “Learning SIMUL8: The Complete Guide” (and SIMUL8 Version 6)", *Journal of Artificial Societies and Social Simulation*, Vol. 6:3, 2003.
- [10] Gulyás, L.: "Generation of Robust Networks with Optimization under Budget Constraints", In *Proceedings of The 5th International Workshop on Emergent Synthesis (IWES'04)*, Budapest, Hungary. 2004.
- [11] Gulyás, L.: "Discrete Choices on Dynamic Networks", Manuscript, *Santa Fe Institute Working Paper*, Santa Fe, NM, USA. To Appear.
- [12] Gulyás, L.: "Design Patterns for Agent-Based Simulation", Submitted to the ‘Artificial Economics’ Symposium in Agent-Based Computational Methods in Finance, Game Theory and their Applications, Lille, France, September 15-16, 2005.
- [13] Gulyás, L. and Adamcsek, B.: "Charting the Market: Fundamental and Chartist Strategies in a Participatory Stock Market Experiment", In *Proceedings of the International Conference Experiments in Economic Sciences: New Approaches to Solving Real-world Problems*, Okayama and Kyoto, Japan, 14-17 December 2004.
- [14] Gulyás, L., Adamcsek, B. and Kiss, Á. "Experimental Economics Meets Agent-Based Finance: A Participatory Artificial Stock Market", In *Proceedings of 34th Annual Conference of International Simulation and Gaming Association (ISAGA 2003)*, To Appear. 2003.
- [15] Gulyás, L., Adamcsek, B. and Kiss, Á.: "An Early Agent-Based Stock Market: Replication and Participation", *Rendiconti Per Gli Studi Economici Quantitativi*, Volume unico, pp. 47-71, 2004.
- [16] Gulyás, L. and Bartha, S.: "How Much Is Too Much? -- What Programming Skills Are Really Needed to do ABM?" In *Proceedings of Seventh Annual Swarm Users/Researchers Conference (SwarmFest 2003)*, Poster, Notre Dame, Indiana, USA. 2003.
- [17] Gulyás, L. and Dugundji, E. R.: "Discrete Choice on Networks: An Agent-Based Approach", In *Proceedings of North American Association for*

- Computational Social and Organizational Science Conference (NAACSOS 2003), June 22-25, 2003, Omni William Penn, Pittsburgh, PA. 2003.*
- [18] Gulyás, L. and Kozsik, T.: "Model Design Interface – A CASE Tool for the Multi-Agent Modeling Language", In *Proceedings of The 2nd International Conference of PhD Students, Section of Engineering Sciences*, Miskolc, Hungary, pp. 97-104., 1999.
 - [19] Gulyás, L. and Kozsik, T.: "The Use of Aspect Oriented Programming in Scientific Simulations", *Software Technology, Fenno-Ugric Symposium FUSST'99 Proceedings* (Ed. Penjam, J.), University of Tallinn, Estonia, pp. 17-28. 1999.
 - [20] Gulyás, L., Kozsik, T. and Corliss, J. B.: "The Multi-Agent Modeling Language and The Model Design Interface", *The Journal of Artificial Societies and Social Simulation*, Vol. 2:3, 1999.
 - [21] Gulyás, L., Kozsik, T., Czabala, P. and Corliss, J. B.: "Telemodeling – Overview of a System", In *Proceedings of Proceedings of the XV. IFIP World Computer Congress '98, 31 Aug - 4 Sep 1998*, Austrian Computer Society (OCG), Vienna/Austria and Budapest/Hungary, pp. 59. 1998.
 - [22] Gulyás, L., Kozsik, T. and Fazekas, S.: "The Multi-Agent Modeling Language", In *Proceedings of Proceedings of the 4th International Conference on Applied Informatics (ICAI)*, Eger-Noszvaj, Hungary, pp. 43-50. 1999.
 - [23] Gulyás, L. and Mansury, Y.: "Patterns of Firm Agglomeration: An Autonomous Agent-Based Model of City Formation", In *Proceedings of International Conference on Complex Systems (ICCS2002)*, Nashua, NH, USA. 2002.
 - [24] Gulyás, L. and Mansury, Y.: "Explaining the Zipf-Law Pattern of City Systems: An Agent-Based Simulation Approach", Manuscript, *Submitted to Journal Of Economic Dynamics and Control*. 2005.
 - [25] Gulyás, L. and Vincze, T.: "Automated Trading Experiments with MAML", *Agent-Based Methods in Economics and Finance: Simulations in Swarm* (Eds. Luna, F. and Perrone, A.), Kluwer Academic Publishers, pp. 39-61. 2002.
 - [26] Mansury, Y. and Gulyás, L.: "Explaining The Zipf-Law Pattern of City Formation: An Agent Based Approach", In *Proceedings of Association of Collegiate Schools of Planning (ACSP) 2002 Conference*, Whashington D.C., USA. 2002.
 - [27] Tatai, G. and Gulyás, L.: "Agents and Multi-Agent Systems", *Artificial Intelligence (textbook)* (Ed. Futó, d. I.), Aula Kiadó, Budapest, Hungary, pp. 709-755. (In Hungarian). 1999.
 - [28] Tatai, G. and Gulyás, L. (eds.): *Agents Everywhere*, Proceedings of the 1st Hungarian National Conference on Agent-Based Computing (HUNABC'98), Springer, Budapest, p. 200. 1999.