On The Transition to Agent-Based Modeling: Implementation Strategies from Variables to Agents

László Gulyás[1]
Government Department
Harvard University

gulyas@fas.harvard.edu

Abstract

This paper studies the effect of different implementation approaches on agent-based computer models. This is accomplished via four re-implementations of a simple model of self-organization. How implementation choices 'guide our hands' and lead possibly to implicit assumptions about the modeled system is also demonstrated. Furthermore, the question of what makes a model agent-based is studied. An argument is made that agent-based implementation is rather a matter of degree than a binary choice.

1.0 Introduction

This paper studies the effect of different implementation approaches on agent-based computer models. While this may seem a meticulously technical issue, it is shown that it may have a dramatic effect on the results. It also demonstrates how implementation choices 'guide our hands' and may lead to implicit assumptions about the modeled system. It is argued that the connection between the conceptual model and its actual implementation is rather important. Especially, as the later rarely gets published.

Another aim of the paper is to explore *what makes a model agent-based*. With the spread of computational social science [1][9] and with agent-based modeling gaining popularity [7] [2], arguments about agent-based versus equation-based modeling (or about individual-based versus variable-based modeling) abound. [11][3][10] In contrast, this paper refrains from advocating a particular approach. Instead, it argues that an agent-based implementation is rather *a matter of degree* than a binary choice.

The study reported here consists of four re-implementations of a simple, but well-known model. The versions range from the original, variable-based formulation to one containing autonomous agents with bounded rationality. Section 2.0 discusses the original model in detail. The framework for the case study is introduced next. This is followed by the description of the different implementation strategies, accompanied by the issues and lessons of the given approach. Section 5.0 concludes the paper.

2.0 Krugman's City Formation Model

The material of this study is a simple, introductory model of self-organization in spatial economics by Paul Krugman. [4] It deals with the emergence of office districts in a polycentric metropolitan area (e.g., like Los Angeles). More precisely, the focus is on the structure that arises from the internal logic of the system, not from inherent

[1] Part of this work has been carried out at the Systems Laboratory, Central European University and at MTA SZTAKI, Budapest, Hungary.

differences among locations. Thus, the model studies the self-organization of space based on independent location decisions by individual firms. [4]

It is assumed that the firms make their relocation decisions depending only on where the other firms are located. The dynamics of the system is driven by the tension of two forces: companies might dislike having other firms nearby (*centrifugal force*[2]), while on the other hand, they might like being close to each other (*centripetal force*). It is assumed that both of these forces decline with distance, but centripetal force declines faster.

The result of this simple setup is that practically *any initial distribution* of business across the landscape, no matter how even (or random), will spontaneously organize itself into a pattern with multiple, clearly separated business centers. Moreover, and more interestingly, these 'edge cities' will be roughly evenly spaced. For example, with the parameter set Krugman describes [5], two equal-sized giant cities emerge out of a flat landscape. In this case, the three-dimensional visualization of the system used in [4] takes a form similar to a '59 Cadillac. [6] (See Figure 2) for an illustration.) This trademark image will be used as the "litmus test" of the replication efforts below.

The selection of this model was motivated by its conceptual clarity and by its easy-to-recognize qualitative outcome. Also, the model was published in a popular and accessible book with information sufficient for replication. Furthermore, the model was explicitly designed to demonstrate self-organization and the bottom-up approach associated with it.

2.1 The Original Formulation

Krugman's results are based on a simplified version of the conceptual model introduced above. It assumes a one-dimensional metropolitan area, which is represented by a circular line (see Figure 1).

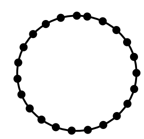


Figure 1

Let x be some location on this circle and let $\lambda(x)$ denote the density of firms at that location. The market potential of a given location depends both positively and negatively on the density of firms at other locations, and both forces decline with distance, according to (1). Here A denotes the strength of the centripetal and B stands for that of the centrifugal force, while r_1 and r_2 are the rates of decline for the two forces, respectively. It is assumed that $r_1>r_2$.

$$P(x) = \int_z \left(A \cdot e^{-r_1|x-y|} - B \cdot e^{-r_2|x-y|} \right) \lambda(z)dz \tag{1}$$

$$\overline{P} = \int_x P(x)\lambda(x)dx \tag{2}$$

$$\frac{d\lambda(x)}{dt} = \gamma\left(P(x) - \overline{P}\right)\lambda(x) \tag{3}$$

$$\lambda_i = \frac{k+u_i}{\sum_j \left(k+u_j\right)} \tag{4}$$

The dynamics of this minimalist model is driven by the assumption that firms gradually migrate toward locations with above average market potential and away

---

[2] Although, the terms *attraction* and *repulsion* may be more appropriate here, the terminology of the original publication is kept.

from those with below average values. Thus, having average potential defined as (2) the dynamic rule takes the form of (3). The initial market share distribution is set according to equation (4), where k is a 'smoothing' parameter and u_i are random values between 0 and 1 drawn from a uniform distribution.

3.0 The Framework of the Case Study

In the previous section Krugman's conceptual model was presented, followed by its mathematical formulation. In a sense the latter is the implementation of the former provided by Krugman. This section turns the attention towards the questions of implementation as a computer simulation.

Outcome *(results)*	Spatial distribution of firms
Information used in rules (*access constraint*)	?
Rules at the level of… *(dynamics)*	?
Implemented Entities *(representation)*	?
Modeled Entities *(consideration)*	Firms

Table 1 Framework for Implementation Strategies

The most elementary issues are what entities or processes are considered in the model, and what outcome do they produce. These questions are answered in the conceptual model: firms are the objects of the study and they are expected to produce agglomerations. However, the dynamics of the system must obviously be expressed in terms of 'objects' present in the implementation. Therefore, the next issue is what entities are actually represented. Krugman's formulation features market shares ($\lambda(x)$) in this role.

In principle, the description of dynamics contains rules that operate on certain entities, based on information extracted from the system. Notice, however, that the represented entities could be different from those of the outcome, thus creating a hierarchy of entities in the implementation. Therefore, it must be decided at which of these levels the rules operate. Also, the way information is gathered for the rules needs to be specified. As the formulation in the previous section only represents entities at the level of its results, it obviously introduces the dynamics at this level, too. Finally, it extracts information about the system in terms of market potential and assumes global access for each entity.

	Krugman's version	**Marionettes**	**Autonomous Agents**	**Bounded Rational Agents**
Information	Global	Global	Global	Local
Rules	Market shares	Market shares	Firms	Firms
Represented	Market shares	Firms	Firms	Firms

Table 2 The Implementation Strategies of the Case Study

With these answers the specification is complete. It is important to note, however, that only the first two answers were derived from the conceptual model. All the rest represent implementation choices. (See Table 1.) The implementation strategies discussed in the next section are chosen by gradually changing the answers given above. (See Table 2.)

The first step is the computational version of the mathematical formulation discussed above. Next, firms are introduced as represented entities, but the rules still operate on market shares. They determine the number of firms to be located at a certain location after each step. In the third step this top-down method is changed, making the firms responsible for their own migratory behavior. In the final step, firms are confined to a certain range of the landscape around themselves to gather information from.

4.0 From Variables To Agents: Four Implementation Strategies

4.1 The Original Formulation

This step implements Krugman's original formulation. The only change required is the discretization of equations (1) and (2). They take the form of (5) and (6), respectively.

$$(5)\ P(x) = \sum_z \left(A \cdot e^{-r_1|x-y|} - B \cdot e^{-r_2|x-y|} \right) \lambda(z) \qquad (6)\ \overline{P} = \sum_x P(x)\lambda(x)$$

Results

No surprising problems occur at this step. Naturally, the technical problems of implementing an analytical model on a computer must be dealt with. Among these is one similar to the issue that led to the concept of 'butterfly effect'. [12] There, Lorenz found that the slightest change in the precision by which his system's initial conditions were given led to dramatic changes in the overall results. In this case, it turns out that equation (3) generates only minute changes in the beginning. Therefore, unless special care is taken the system is stuck in its original state.

Normally, however, simulation reveals that this implementation replicates the qualitative behavior reported in the original publication. (See Figure 2) and [4].)[3] Figure 2) shows the distribution of market shares (vertical axis) among locations (the axis leaning to the right) in simulation time (left-leaning axis). Clearly, two evenly spaced giants emerge out of a fairly even distribution.

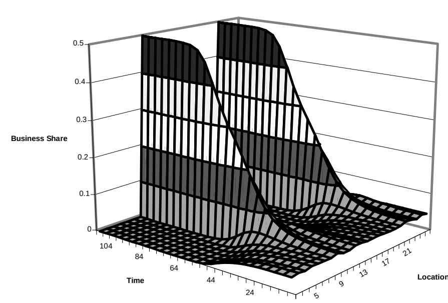


Figure 2 Results of the Original Formulation

Quantitatively, however, there is a difference as Krugman's cities emerge in 25 iterations, while the replica produces them over a significantly longer time period. This discrepancy is suspected to be a consequence of a combination of rounding issues and the possibly different precision of the computer arithmetic used.

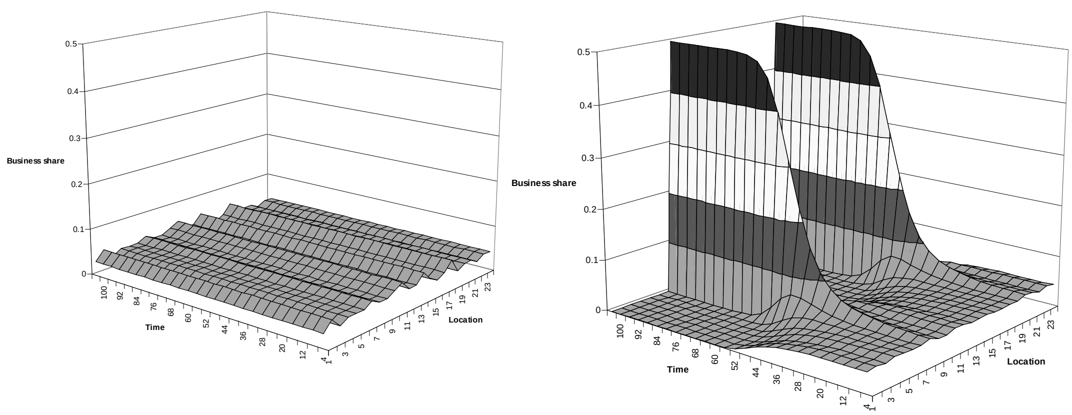
[3] Note that the model is deterministic except the initial distribution of densities. Therefore, the result of a single run is presented here.

4.2 'Marionettes'

The previous implementation featured no representation for the firms. In this step explicit representation of the model's agents is introduced, preserving the way the dynamics of the system is formulated. Thus, firms are implemented as agents without autonomous control, i.e., 'marionettes', and migratory behavior is guided by system level aggregate properties.

Technically, market shares are determined according to the actual allocation of firm 'objects'. From this the distribution of market potential follows by (5). Then, for each location, the number of agents to leave is calculated based on (3).

In principle, the rest is as simple as picking the ones to leave and move them to their new location. However, there are two problems. First, all agents at the same place are identical. Therefore, random selection among them is applied. Second, the model does not specify where exactly to move these agents. The only specification is that they move to one of the locations where the density change is positive. In this implementation the following technique is used. The selected 'movers' from each location with a net loss are placed in an 'agent pool'. Then, all locations with a net gain are assigned the appropriate number of new residents from the pool.



a) Moderate-sized population (100) **b) Large population (10000)**

Figure 3 Results of the 'Marionettes' Version[4]

Results

The main lesson is that by the introduction of an entity-level representation of firms the problem is discretized once again. This creates a granularity issue, since fractions of firms are meaningless and only whole agents can exhibit behavior. Therefore, a new parameter, the number of firms gains importance. As migration is expressed as a percentage of the population, the dynamics of the system strongly depends on the size of it. In the early stages of the simulation migration is minute yet, so it requires a population of a certain size to make it manifest itself. Without that the system freezes into its original state. Figure 3a) demonstrates this effect by showing the evolution of a system with 100 agents only. In contrast, the size of the population featured on Figure 3b) is 10.000, which clearly replicates the results of Krugman. This

[4] This step, such as the ones following it, deals with a stochastic model. Therefore, discussion of results should be based on a series of simulation runs. This is omitted here, since the qualitative behavior reported here does not depend on the particular random number sequence.

dependence was not anticipated in the conceptual model. Thus, it may be interpreted as a hidden assumption.

Another lesson was that system level aggregate values, describing the firms' moving behavior, do not account for all actual moves. Rather, they only contain the vector sum of individual actions. This means that individual behavior cannot be unambiguously derived from the conceptual rules of dynamics.

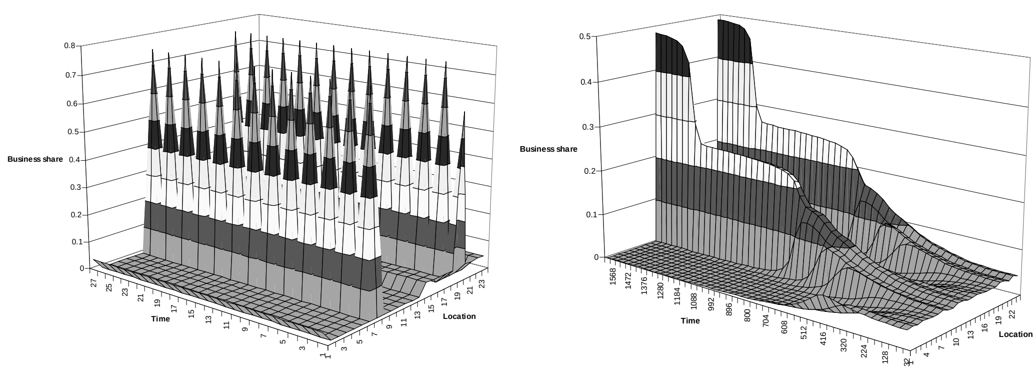
A related problem is that agents at the same location are indistinguishable. Therefore, the conceptual model provides no information on exactly which agents move. The implementation discussed here makes a random selection among the co-located agents.

4.3 Let the Agents Live Their Life: Autonomous Agents

This step of the study introduces autonomous agents. For this, agents need control over their movements. Therefore, the system dynamics is rebuilt from the bottom up, replacing the top-down logic of equation (3) with a stochastic approach.

First, market shares and potentials are calculated the same way as in the previous step. Then each agent gets the chance to move. However, modeled on equation (3), the probability that an agent actually moves is proportional to the amount by which its current location's potential is below the average. Since, negative differences are mapped to zero probability, no agent can leave an above-average location.

If an agent does move, it does that, too, randomly. It assigns probabilities to above-average locations in proportion to their potential exceeding the average. Then it selects its next office based on that distribution.



a) Maximizing agents (N=10000) **b) The effect of randomization (N=50000)**

Figure 4 Results of the Autonomous Agents Implementation

Results

It is important to realize that neither the conceptual model, nor the original implementation provide any information about the agent-level behavior of the firms. Therefore, the number of possibilities to fill in this gap is limitless. In this study, a stochastic approach was used, partly for the sake of simplicity.

On the other hand, Thomas and Sycara suggest that heterogeneity may be essential to achieve equilibrium. [8] Indeed, Figure 4a) shows what happens if firms choose their best alternatives, i.e., the location with the highest market potential. Clearly, a giant agglomeration emerges promptly, but then it keeps 'jumping' endlessly between two sites on the landscape.

Randomization is a straightforward way to introduce heterogeneity and takes care of this problem as shown on Figure 4b). The figure, however, also displays an increased time-line. This is due to another dependence on the population size. Technically, slight changes generated by equation (3) are mapped into very small probabilities in this implementation. Thus, the more agents take a chance against that probability, the more likely the associated event becomes.

4.4 Bounded Rational Agents: Localizing the Control

The implementation strategies above all make a rather strong assumption. They all contain omniscient agents. This assumption corresponds to the notion of rationality, even though the stochastic implementation already weakened it by a little. This final step of the study imposes bounds on the firms' rationality by limiting the information they can collect from the world they live in.

The underlying logic of this implementation strategy is the same as the one in the previous section, except that firms can only see to a certain distance. The range of this 'vision' is a new parameter and takes the same value for all agents. Market shares and potentials are calculated according to the same principle as above, but are now based on this localized information. Therefore, perceived market shares and potentials differ at different locations. Also, the migratory behavior corresponds to the one in Section 4.3, except that it works with the locally perceived potentials.[5]

It is worth noting that the Autonomous Agent implementation above is a special case of this one with the range of vision corresponding to the whole landscape.

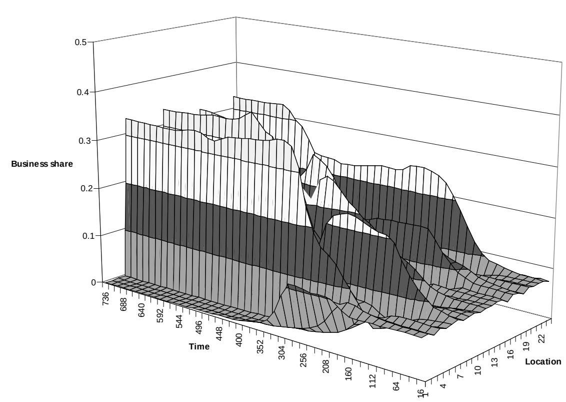


Figure 5 The World of Bounded Rational Agents

Results

Limited information is a radical step away from Krugman's original formulation. Still, the principle, that the model leads to self-organization, i.e., that evenly spaced agglomerations emerge from an even landscape, holds. Figure 5) shows four cities

---

[5] Notice that this prevents firms from directly moving to locations outside their range of vision. This corresponds to the principle that firms behave rationally, subject to the limitations imposed on them.

forming. Obviously, the number of agglomerations depends on the ratio of the firms' vision and the size of the world. In the case of the figure, the agents have information about the half of the landscape surrounding them.

The implementation strategy discussed in this section adds another level of heterogeneity to the system. It introduces localized knowledge. Also, it demonstrates that bounded rational agents can reproduce the basic insight of a conceptual model, even if the original formulation used a top-down approach.

5.0 Conclusions

This paper reported on a case study consisting of four re-implementations of the same conceptual model. The point was to demonstrate that agent-based conceptual model might imply radically different actual implementations. Also, it was argued that there is a continuum of choices available, not only the two idealized extremes of agent-based and equation-based formulation. This suggests that arguments over these two genres of modeling may be strongly improved by taking this insight into account.

The choice of implementation strategy is a technical issue. Nonetheless, it was demonstrated that it plays an important role in determining the model's outcome. Therefore, it should follow interaction topology and activation regime on the imaginary list of important design choices in computational social science. [13][14].

6.0 References

[1] Gilbert G. N., Conte R. (eds.): *Artificial Societies: The computer simulation of social life*, London, UCL Press, 1995.

[2] Gulyás L., Kozsik T., Corliss J. B.: *The Multi-Agent Modeling Language and The Model Design Interface*, Journal of Artificial Societies and Social Simulation vol. 2, no. 3, http://www.soc.surrey.ac.uk/JASSS/2/3/8.html

[3] Parunak H. V-D., Savit, R., Riolo R. L.: *Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide*, In the Proceedings of the Workshop on Modeling Agent Based Systems (MABS98), Paris, 1998.

[4] Krugman, P.: *The Self-Organizing Economy*, Blackwell Publishers Inc., 1996, pp. 101-106.

[5] Same publication, page 106.

[6] Same publication, page 26.

[7] Minar N., Burkhart R., Langton C., Askenazi M.: *The Swarm Simulation System: A Toolkit for Building Multi-agent Simulations*, http://www.santafe.edu/projects/swarm/overview/overview.html

[8] Thomas, J. D., Sycara, K.: *Heterogeneity, Stability, and Efficiency in Distributed Systems*, In the Proceedings of the 1998 International Conference on Multi-Agent Systems, July 1998, pp. 293-300.

[9] Axelrod, R.: *Advancing the Art of Simulation is Social Sciences*, In *Simulating Social Phenomena*, Conte, R., Hegselmann, R. and Terna, P. (eds.), LNEMS-456, Springer, 1997. pp. 21-40.

[10] Brassel, K-H., Möhring, M., Schumacher, E., Troitzsch, K. G.: *Can Agents Cover All the World?*, In *Simulating Social Phenomena*, Conte, R., Hegselmann, R. and Terna, P. (eds.), LNEMS-456, Springer, 1997. pp. 55-72.

[11] Epstein, J. M., Axtell, R.: *Growing Artificial Societies: Social Science from the Bottom Up*, The MIT Press, 1996.

[12] Lorenz, E. N.: *Predictability: Does the Flap of a Butterfly's Wing in Brazil Set off a Tornado in Texas?*, Talk at the meeting of the American Association for

the Advancament of Science, Washington, D.C., 1972. Also in *The Essence of Chaos* (Lorenz, E. N.), University of Washington Press, 1993.

[13] Axtell, R.: *Effects of Interaction Topology and Activation Regime in Several Multi-Agent Systems*, SFI Working Paper 00-07-039, 2000.

[14] Cohen, M. D., Riolo R. L. and Axelrod, R.: *The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and Other Factors Promote Cooperation*, SFI Working Paper 99-01-002, 1999.