**Cellular automata and agent-based models: what next?**

Helen Couclelis

Cellular automata (CA) and agent-based models (ABM) are hallmarks of computational geography. Increasingly they are used in combination in the development of process-oriented models of people or other kinds of actors interacting with their environment, with CA typically simulating the spatial environment and ABM representing the relevant decision-making units. Combined CA-ABM models may be used to simulate farmer communities dynamically affecting land cover through adaptive land use decisions, households seeking suitable housing in changing urban areas, or trip makers responding to various congestion pricing schemes on transportation networks. Such models are built to incorporate considerable empirical and intuitive understanding of the complex processes of interest, and when calibrated to actual data they are often presented as suitable for prediction and policy analysis. As someone who has studied these two types of models for over 20 years, I am skeptical of such claims (Couclelis 2001). I believe that we have by now accumulated enough experience for a more systematic exploration of the potential and the limitations of CA and ABM to model complex spatial systems, whether used in conjunction or separately.

Computation has often been discussed as the third way of doing science, lying somewhere between theory development and experimentation. This implies a new approach to knowledge production and the need for a new kind of research methodology different from either the mostly deductive mode of theoretical work or the mostly inductive mode of experimental science. That third way centers on the construction of complex simulated worlds within which experiments may be run that would have been difficult or impossible to conduct in the real world. The epistemological problem is that models of complex open systems with deep uncertainties, as social systems nearly always (and natural systems usually) are, cannot in principle be used for prediction. Predictive models belong in the traditional scientific paradigm of theoretical closed system descriptions supported by experimental evidence, or at least of well-established empirical generalizations such as human geography's spatial interaction models. Because this fact is not always appreciated, many computational modelers understand progress in the field to mean building simulations that are increasingly detailed and realistic, even though increased detail can actually decrease any predictive value such models may have. The meaning of model validation in this new world of computational process models thus remains open, and so does the question of how to derive valid insights that may be useful for both theory development and for policy guidance.

Technically, the reason why models of complex open systems cannot yield reliable predictions is that many (in some cases infinitely many) different models can provide acceptable fits to the data. In other words, any particular model is but one realization out of a large space of potential models, few or none of which may be correct by whatever definition of the term. This issue is sometimes addressed with Monte Carlo simulations that generate many versions of a particular model by systematically varying the parameters; model outcomes are then considered reliable to the extent that they are reproduced by large numbers of different parameter sets. This methodology may take care of parametric uncertainty but cannot address structural uncertainty, that is, the degree of confidence one may have in the structural validity of the model. Researchers in both the social and the natural sciences have suggested methods for generating large numbers of different model structures in a manner analogous to generating versions of the same model through Monte Carlo simulations. The idea is that investigating the properties of entire ensembles of models, even relatively simple ones, may yield more robust insights into the complex spatial processes of interest than the study of even the more realistic-looking individual models. Procedures for generating ensembles of models for that purpose have been described in hydrology by Beven and associates in a long series of papers (e.g, Beven and Freer 2001), in policy studies by the RAND team of Popper, Lempert and Bankes (e.g., Popper et al. 2003), and in several other fields.

Should we wish to explore that direction, our task will be greatly facilitated by the fact that in formal terms, CA and ABM are very close cousins. Both are structures described in the theory of automata, one of the three major branches of the mathematical theory of computation. A CA may be seen as spatial array of ABM. In principle, anything that can be modeled as a (generalized) CA can also be modeled as an ABM and vice versa, though obviously some options will be more intuitive and computationally efficient than others. Thus CA models have been developed where the cells are endowed with complex goal-directed decision rules and ABM where the agent is the environment. Some researchers consider mobility to be the defining difference between the two kinds of models, but in actual fact CA simulate movement in the same way your computer screen does, by spreading activation down a sequence of adjacent cells or pixels. (Action at a distance – easy for ABM – is somewhat trickier to simulate within a pure CA framework, but that too can be done). The affinity between ABM and CA means that both agents and environment can be specified within the same framework in the formal language of automata theory. Such integration is very likely to provide a substantially increased theoretical understanding of the properties of these structures and to greatly support the generation and analysis of appropriate ensembles of models. I think that there is fertile ground here for the more theoretically inclined among us to make contributions to complex spatial systems modeling that could benefit researchers in many fields.

**References**

Couclelis, H. (2001). Why I no longer work with agents: a challenge for ABMs of human-environment interaction. In D. Parker, T. Berger, and St. Manson (Eds.) *Agent-based models of land-use and land-cover change*, LUCC Report Series No.6, 3-5.

Beven, K. and J. Freer (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249 (2001), 11-29.

Lempert, R.J., S. Popper and S. Bankes (2003). *Shaping the next one hundred years: new methods for quantitative long-term policy analysis.* RAND, Santa Monica, CA.