University of Groningen

Social mechanisms and generative explanations

Macy, M.W.; Centola, D.; Flache, A.; van de Rijt, A.; Willer, R.

Published in:
Analytical Sociology and Social Mechanisms

DOI:
[10.1017/CBO9780511921315.013](https://doi.org/10.1017/CBO9780511921315.013)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2011

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Macy, M. W., Centola, D., Flache, A., van de Rijt, A., & Willer, R. (2011). Social mechanisms and generative explanations: Computational models with double agents. In P. Demeulenaere (Ed.), *Analytical Sociology and Social Mechanisms* (pp. 250-265). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511921315.013>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

12 Social mechanisms and generative explanations: computational models with double agents

Michael W. Macy with Damon Centola, Andreas Flache, Arnout van de Rijt and Robb Willer

Traditionally, sociologists have tried to understand social life as a structured system of institutions and norms that shape individual behavior from the top down. In contrast, a new breed of social modelers suspect that much of social life emerges from the bottom up, more like improvisational jazz than a symphony orchestra. People do not simply play parts written by elites and directed by managers. We make up our parts on the fly. But if everyone is flying by the seat of their pants, how is social order possible? The puzzle is compounded by scale. Coordination and cooperation are relatively easy to attain in a jazz quartet, but imagine an ensemble with millions of musicians, each paying attention only to those in their immediate vicinity. Without a Leviathan holding the baton, what prevents the population from descending into a Hobbesian cacophony of all against all?

Social life from the bottom up

New and compelling answers to this question are being uncovered by social theorists using an innovative modeling tool developed in computer science and applied with impressive success in disciplines ranging from biology to physics – agent-based computational (ABC) modeling.

ABC models are *agent-based* because they take as a theoretical starting point a model of the autonomous yet interdependent individual units (the “agents”) that constitute a dynamic system. The models are *computational* because the individual agents and their behavioral rules

This chapter builds on and extends Flache and Macy (2009), Macy and Willer (2002), Centola and Macy (2005) and Macy and van de Rijt (2006) and Van de Rijt, A., D. Siegel and M. Macy (2009). Co-authors are listed alphabetically. The authors wish to thank the National Science Foundation (SBR 0241657 and SES-0432917) and the Netherlands Organization for Scientific Research (NWO-VIDI-452-04-351) for support during the time that this research was conducted.

are formally represented and encoded in a computer program such that the dynamics of the model can be deduced using step-by-step computation from given starting conditions.

ABC modeling was originally developed in computer science and artificial intelligence as a technology to solve complex information processing problems based on autonomous software units. These units can each perform their own computations and have their own local knowledge, but they exchange information with each other and react to input from other agents. ABC models have been used to understand how spontaneous coordination can arise in domains as diverse as computer networks, bird flocks and chemical oscillators.

Increasingly, social scientists are using this same methodology to better understand the complexities of social life as well. Despite their technical origin, agents are inherently social. Agents have both a cognitive and a social architecture (Gilbert and Troitzsch 1999; Wooldridge and Jennings 1995). Cognitively, agents are heuristic and adaptive. Socially, agents are autonomous, interdependent, heterogeneous and relational. *Heuristic* means that agents follow simple behavioral rules, not unlike those that guide much of human behavior, such as habits, rituals, routines, norms, and the like (Simon 1982). *Adaptive* means that actions have consequences that alter the probability that the action will recur, as agents respond to feedback from their environment through learning and evolution. *Autonomous* agents have control over their own goals, behaviors and internal states and take initiative to change aspects of their environment in order to attain those goals. They interact with little or no central authority or direction. However, autonomy is constrained by behavioral and strategic *interdependence*. Behavioral interdependence means agents influence others in response to the influence that they receive, through persuasion, sanctioning and imitation. Strategic interdependence means the consequences of each agent's decisions depend in part on the choices of others. *Heterogeneity* relaxes the assumption in system dynamics models (and many game theoretic models as well) that populations are composed of representative agents.

Finally, and perhaps most importantly, agents are relational; that is, they interact locally with "neighbors," such that population dynamics are an emergent property of local interaction. ABC models allow researchers to study how local decision heuristics and network topology interact to produce highly complex and often surprising global patterns.

Can social scientists learn something from models developed for understanding computer networks, bird flocks or chemical oscillators? Agent modelers believe they can, for several reasons. First, ABC models show how very simple rules of local interaction can generate

highly complex population dynamics that would be extremely difficult (if not impossible) to model using traditional methods. Second, these models show how Durkheimian “social facts” can emerge *sui generis* at the population level, even when these properties do not exist at the level of the interacting agents. Third, these models can be used as virtual laboratories, to reveal the micro mechanisms responsible for highly complex social phenomena.

The microfoundations of social complexity

Like game theory, ABC modeling is a formal implementation of “methodological individualism,” the search for the microfoundations of social life in the actions of intentional agents. This is the major difference between ABC modeling and an earlier generation of equation-based methods of computer simulation such as system dynamics. Traditional computer simulation modeled population behavior as a unified system. These models generated population patterns based on interactions among system-level processes, usually expressed as rates of change among system components. ABC models generate population patterns as an emergent property of local interactions among autonomous decision-makers. Instead of a model of the population, we have a population of interdependent models, each corresponding to an individual actor. Contrary to conventional intuitions, ABC models have repeatedly demonstrated how the emergence of macro-social patterns out of micro-social interactions is not always a matter of simple aggregation.

ABC modeling also differs from game theory in relaxing the behavioral and structural assumptions required for mathematical tractability. The attraction to ABC modeling centers on the ability to use computation to solve problems that would otherwise be mathematically intractable, such as systems with nonlinear dynamics, complex networks and sensitive dependence on initial conditions. As in the physical and life sciences, social scientists have begun to appreciate that the complexity of social systems cannot always be understood using traditional deductive methods.

Thomas Schelling’s (1971) canonical tipping model of residential segregation was among the first efforts to extend game theory by letting the players walk through the problem rather than trying to solve the problem mathematically. Consider a city that is highly segregated, such that every resident has neighbors with identical cultural markers (such as ethnicity). If the aggregate pattern were assumed to reflect the attitudes of the residents, one might conclude from the ethnic distribution that the population was highly parochial and intolerant of diversity. Yet Schelling’s tipping model shows that this need not be the case. He assumes that

residents are tolerant of ethnic diversity but do not want to become a minority in the neighborhood. The model shows that perfectly segregated neighborhoods will form even in a population that is tolerant of diversity. The aggregate pattern should therefore not be taken to reflect the underlying attitudes of the constituent individuals. The first agent who moves out to avoid being in the minority makes the imbalance even worse for the agent's co-ethnic neighbors. This ethnic flight, in turn, adds another majority-group member to the receiving neighborhood, thereby making the ethnic distribution even less tolerable for the minority members.

A computational extension of the Schelling model

Schelling's classic model did not require a computer. The model was implemented on a large checkerboard using red and blue poker chips. This implementation made it necessary to keep the model extremely simple. As a consequence, the model imposed three simplifying assumptions that can be relaxed by implementing the model on a computer. First, Schelling assumed that agents are tolerant of diversity but nevertheless prefer co-ethnic neighbors. It is not all that surprising that segregation would emerge in a population in which everyone is ethnocentric, even if they tolerate diversity. We do not need computer simulation to immediately recognize that total segregation is an equilibrium in this population. Suppose instead that agents have multiculturalist preferences for diversity, such that they would choose a mixed neighborhood over one that was entirely co-ethnic. Would this prevent the population from tipping into segregation? Second, Schelling assumed that decisions about where to live were deterministic. Suppose there is a small amount of noise in decision-making. Might this destabilize a segregated equilibrium? Third, Schelling assumed that agents were indifferent to ethnic composition on either side of a single threshold. Suppose the probability of moving increases continuously with the discrepancy between the observed and desired ethnic composition of the neighborhood. Put differently, suppose the agents are highly sensitive to even very small changes in the ethnic composition of their neighborhood.

Van de Rijt *et al.* (2009) relaxed these three simplifying assumptions by implementing the preference function used in a mathematical model developed by Zhang (2004; see also Panks and Vriend 2007), in which agents strongly prefer neighborhoods with a mixed composition that reflects the population distribution of ethnicity. In addition, agents' moves are stochastic, based on a probability distribution that corresponds to their preferences. Agents are thus least likely to move out when the current neighborhood reflects the ethnic diversity of the population and most likely to move out when the local population is homogenous,

1) Low Noise

2) High Noise

The discontinuous line indicates a discontinuous (threshold) preference for diversity. The solid line indicates a linear preference. These preferences are weaker relative to chance in panel 2.

Figure 12.1 Multicultural preferences.

but slightly less so if this homogeneity favors their own ethnic group. Finally, in one condition, agents are sensitive to small changes in ethnic composition, and in the other condition, agents have a threshold level of ethnic homogeneity above which they move out in search of a neighborhood with greater diversity, as depicted in Figure 12.1.

When the noise level was set to zero and the preference function was linear, the results confirmed Zhang's (2004) mathematical results – the population segregated. Moreover, this outcome was highly robust to the introduction of a small amount of noise in the decision function. However, when the linear function was replaced with a threshold function, the population remained integrated. A multiculturalist population with a threshold preference for diversity is less likely to tip into segregation compared to a similarly multiculturalist population that is sensitive to small changes in the ethnic composition of their neighborhood.¹

The mechanism that allows a multicultural population to segregate is the asymmetry (evident in Figure 12.1) in dissatisfaction with neighborhoods that are ethnically imbalanced. This dissatisfaction is slightly

¹ Using a computational model in which agents prefer as many co-ethnic neighbors as possible, Bruch and Mare (2006) reached the opposite conclusion – that tipping at the individual level (i.e. a threshold for moving out) leads to tipping at the population level (i.e. a cascade leading to segregation). However, van de Rijdt *et al.* (2009) showed that this result was caused by an error in their model, and that the integration they observed with linear preferences was possible only when ethnic preferences were sufficiently weak relative to chance, an error Bruch and Mare acknowledged (2009) in response.

weaker when the imbalance favors the ingroup than when an identical imbalance favors the outgroup. Although this difference in dissatisfaction is very small compared to the much larger difference in the preference for balanced vs. imbalanced neighborhoods, the small ingroup bias is sufficient to tip the population into a highly segregated pattern when agents are sensitive to small changes in ethnic composition. Even in a population that is highly integrated, there will be some neighborhoods that will not be perfectly balanced – one of the two groups will be slightly favored. In these minimally imbalanced neighborhoods, agents will have a very low probability of moving out, but the probability will nevertheless be slightly higher if the imbalance favors the outgroup. Over time, this small bias inevitably causes the level of imbalance to increase, which in turn confronts agents with an ever-starker choice between neighborhoods that favor – by an ever-widening margin – one group over the other. The only possible outcome is then a maximally segregated population that leaves everyone dissatisfied except the lucky few who manage to end up on the borders between the ghettos.

However, agents with threshold preferences manage to remain integrated despite a comparable asymmetry in their dissatisfaction with imbalanced neighborhoods. That is because the asymmetry only applies to agents who live in neighborhoods that exceed their tolerance level for ethnic imbalance. As long as the level of imbalance remains tolerable, these agents do not care whether the imbalance favors their own group or the other. And as long as the agents do not care, the level of imbalance remains within their tolerance level.

In short, these results show that Schelling actually understated the possibility for segregation to emerge in a population that does not seek segregation. He showed how this can happen in a population that tolerates ethnic diversity up to a point. We now see how segregation can emerge even in a population that prefers integration, so long as the residents are sufficiently sensitive to small changes in ethnic composition.

The double edge of networks

An ABC model of cooperation among friends also showed how individual behaviors can aggregate with opposite effects on the behavior of the population. In network analysis, an “edge” is an undirected tie between two nodes, such as two friends or two neighbors. Flache and Macy (1996) discovered that friendship networks can have a “double edge,” meaning that the tie between two friends can have opposite effects on the solidarity of a group to which both friends belong. On the one side, people usually care more about approval from their friends

than from strangers, and this can promote cooperative behavior that is enforced informally by peer pressure to conform to group obligations, particularly in groups whose members are highly dependent on one another for social approval (such as residents of a small town with limited opportunities for outside friendships). It can be proven mathematically that it is an equilibrium when everyone contributes and everyone approves of their friends.² Moreover, the greater the dependence on friends for approval, the stronger the effect of peer pressure in reinforcing this equilibrium. Yet when Flache and Macy tested the effects of dependence on the group using an ABC model of contribution to a public good, the result was surprising: the more dependent people were on approval from their friends, the less willing they were to cooperate. What went wrong? Although dependence on friends for social approval made informal sanctions more effective, it also made friends reluctant to risk friendship by criticizing one another. The mechanism is structural. Because friendship is dyadic, coordination on mutual approval is far more likely than the multilateral coordination required for the exchange of contribution for approval. As a consequence, agents exchanged approval directly for approval, even when both friends failed to contribute to the group.

Contrary to intuition, a stronger need for approval at the micro level did not lead to more effective social control at the macro level.

An agent-based model of naked emperors

Centola *et al.* (2005) discovered another population dynamic with surprising microfoundations. They modeled the curious tendency for people to pretend to believe something they know not to be true, and to disparage those who disagree, a behavioral pattern that is portrayed in Hans Christian Andersen's story of the *Emperor's New Clothes*. Everyday examples include college students who celebrate and encourage intoxication through drinking games, yet who privately express discomfort with excessive consumption (Prentice and Miller 1993). The foible is not limited to students. We all know "naked" scholars whose incomprehensible writings are applauded by those who pretend to understand and appreciate every word, and who disparage skeptics for being intellectually shallow. The pattern extends to closet deviants who publicly disparage others so as to avoid suspicion, such as anxious citizens in totalitarian regimes who denounce their neighbors so as to affirm their loyalty to the regime (Kuran 1995).

² For a comparison of computational and mathematical solutions, see Flache *et al.* (2000).

Centola *et al.* wanted to know if a cascade of compliance could trap a population into thinking a norm was highly popular, when in fact almost everyone shared a strong (but private) wish that the norm would go away. In addition, they examined the structural conditions that might make a population highly vulnerable to such a cascade. Their model assumes a population of agents who are willing to comply with and enforce a norm they privately oppose, but only if the pressure to comply is sufficiently strong. These agents live in a clustered network, but with some overlap between the clusters. The model is initialized with a population that contains only a tiny fraction of vigilant “true believers” who truly support the norm and will always sanction deviant neighbors. The rest of the population opposes the norm, and in the absence of enforcement, the skeptics refuse to comply.

The model demonstrated that widespread enforcement of an unpopular norm can be an equilibrium in a population where the fear of exposure as an imposter is sufficient to motivate enforcement as a way to signal the sincerity of compliance. Yet this equilibrium cannot be reached from a random start in a fully connected population. However, if interaction is restricted to the local neighborhood, and these neighborhoods sufficiently overlap, the equilibrium can be reached, even with only a tiny fraction of “true believers.”

The computational model also identified a surprising condition that is necessary for these cascades to succeed. Contrary to an extensive literature on small worlds (Watts and Strogatz 1998) and “the strength of weak ties” (Granovetter 1973), long-range ties that bridge across clusters actually inhibit the propagation of unpopular norms compared to networks of equal size and density but with fewer ties between clusters. Why this sensitivity to network topology? “Shortcuts” across the network allowed agents in conformist neighborhoods to see deviants in other neighborhoods, which in turn reinforced their private opposition to the norm.

These results deepen our understanding of the dynamics of a puzzling social phenomenon that has perplexed social scientists for years. Further, the results motivated a follow-up laboratory experiment designed to study the false enforcement of an unpopular norm (Willer *et al.* 2009). In the face of social pressure to conform, participants in the experiment not only agreed to a belief they knew to be false, but also publicly penalized a deviant who stated what the participants privately knew to be true. Yet when sanctioning was private, participants favored the deviant. In a related paper, Centola and Macy (2007) investigated in greater detail why shortcuts across a network can inhibit rather than facilitate the propagation of a social contagion, such as conformity to an emergent but unpopular norm. A series of computational experiments highlighted the importance

of a qualitative distinction between “simple contagions” with an “infection” threshold of one (such as disease or information) and “complex contagions” with any threshold higher than one (such as the willingness to act on information). Social contagions are typically complex, due to the social and material costs faced by early adopters of risky innovations, controversial practices, or new technologies whose benefits require widespread usage (e.g. email). As a result, the probability of adoption increases with the proportion of one’s neighbors who have already adopted. This points to the need for an important caveat for Granovetter’s (1973) theory of the “strength of weak ties” that bridge between network clusters. Although clustered sources of information tend to be redundant, clustered sources of advice, persuasion or legitimation are essential for the spread of many social contagions.

The virtues and vices of ABC modeling: the problem of double agents

Agent-based simulations are often criticized on three counts: artificiality, triviality and reliability:

1. Computational models, such as mathematical models, tell us about a highly stylized and abstract world, not about the world in which we live. In particular, agent-based models rely on highly simplified models of rule-based human behavior that neglect the complexity of human behavior.
2. Computational models cannot tell us anything that we do not already know, given the assumptions built into the model and the inability of a computer to do anything other than execute the instructions given to it by the program.
3. Unlike mathematical models, simulations are numerical and therefore cannot establish lawful regularities or generalizations.

1. *Realism.* The defense against these criticisms centers on the principles of complexity and emergence. In complex systems, very simple rules of interaction can produce highly complex global patterns. Although this principle has been mainly applied to physical and biological systems, Simon (1998: 53) draws out the striking implication for the study of human behavior – the possibility that “the apparent complexity of our behavior is largely a reflection of the complexity of the environment,” including the complexity of interactions among strategically interdependent, adaptive agents. The simulation of artificial worlds allows us to explore the complexity of the social environment by removing the cognitive complexity (and idiosyncrasy) of constituent individuals. “Human beings,” Simon contends, “viewed as behaving systems, are

quite simple.” We follow rules, in the form of norms, conventions, protocols, moral and social habits, and heuristics. Although the rules may be quite simple, they can produce global patterns that may not be at all obvious and are very difficult to understand.

In short, the artificiality of agent-based models is a virtue, not a vice. Agent-based modeling does not necessarily “aim to provide an accurate representation of a particular empirical application” (Axelrod 1997: 5). Rather, the goal “is to enrich our understanding of fundamental processes that may appear in a variety of applications.” When computational models are used for simulation or to make predictions or for training personnel (e.g. flight simulators), the assumptions need to be highly realistic, which usually means they will also be highly complicated (Axelrod 1997: 5). “But if the goal is to deepen our understanding of some fundamental process,” Axelrod continues, “then simplicity of the assumptions is important and realistic representation of all the details of a particular setting is not.”

There are several advantages of artificiality. First, lawful regularities may be obscured by conditions in the empirical world that block or distort their expression. Holland (1995: 146) offers a classic example: Aristotle’s mistaken conclusion that all bodies come to rest, based on observation of a world that happens to be infested with friction. Had these observations been made in a frictionless world, Aristotle would have come to the same conclusion reached by Newton and formulated as the principle that bodies in motion persist in that motion unless perturbed. Newton avoided Aristotle’s error by studying what happens in an artificial world in which friction had been assumed away. Ironically, “Aristotle’s model, though closer to everyday observations, clouded studies of the natural world for almost two millennia” (Holland 1995: 146). In contrast, Newton’s model generates more accurate predictions than does Aristotle’s, once the effects of friction are reconsidered. More generally, it is often necessary to step back from the world we experience, into one that is more abstract, in order to see the empirical world more clearly.

For example, Schelling’s neighborhood-segregation model made no effort to be “realistic” or to resemble an observed city. Rather, the “residents” live in a highly abstract cellular world without traffic, poverty, crime, lousy schools, train tracks, zoning laws, red-lining, housing prices, etc. This artificial world shows that all neighborhoods have an underlying tendency toward segregation even when residents are tolerant of diversity, so long as residents do not want to become a local minority. This hypothesis can then be tested in observed neighborhoods. Of course, real world observations may not confirm the model predictions, but this does not detract from the value of the experiment. On the contrary, empirical disparities lead the investigator to look for

conditions that were not present in the model that might account for the discrepancy with the observed outcome.

In short, agent-based models “have a role similar to mathematical theory, shearing away detail and illuminating crucial features in a rigorous context” (Holland 1995: 100). Nevertheless, Holland reminds us that we must be careful. Unlike laboratory experiments, computational “thought experiments” are not constrained by physical reality and thus “can be as fanciful as desired or accidentally permitted. Caution and insight are the watchwords if the computer-based model is to be helpful” (Holland 1995: 96).

2. *The results are wired in.* The second criticism centers on the inability of computers to act in any way other than how they were programmed to behave. This criticism overlooks the need to “determine the logical consistency of a set of propositions and the extent to which theoretical conclusions actually follow from the underlying assumptions” (Prietula *et al.* 1998: xv). Indeed, a properly designed computational experiment can reveal highly counter-intuitive and surprising results. This theoretical possibility corresponds to the principle of *emergence* in complex systems. Biological examples of emergence include life that emerges from non-living organic compounds and intelligence and consciousness that emerge from dense networks of very simple switch-like neurons. Schelling’s (1971) neighborhood model provides a compelling example from social life: segregation is an emergent property of social interaction that does not necessarily reflect the simple aggregation of individual preferences. And Centola *et al.* (2005) show that norms that enjoy widespread public support do not necessarily reflect the distribution of private beliefs. In short, ABC models can be extremely useful in identifying surprising macro-level implications of a set of micro-level assumptions. That these implications are “wired in” does not in any way detract from the value of the computational experiments. On the contrary, by revealing unexpected population patterns that are “wired in” to a set of behavioral and structural assumptions, ABC models demonstrate the danger of trusting “intuition” and “common sense” instead of formal modeling.

Nevertheless, the problem of wiring in the results can also be very serious: when the computational results are sensitively dependent on hidden assumptions that are theoretically arbitrary. Sensitivity to theoretically motivated assumptions is not a problem – on the contrary, that is the reason for running computational experiments in which these assumptions are carefully manipulated. However, results that are an artifact of hidden assumptions that do not correspond to the theory that motivated the research are completely worthless and misleading. This criticism points to the need to make all assumptions explicit and

to test the robustness of the results to alternative specification of those assumptions that are theoretically arbitrary.

3. *Computational models are not deductive.* A third criticism is that computational models, unlike mathematical models, are numerical, not deductive. This has two important implications. First, unlike deductive models that do not depend on particular numerical inputs, ABC models generate population patterns that vary across the parameter space for reasons that can be very difficult to identify, especially if the models are overly complicated. Agent models can reveal correlations between inputs and outputs even if the modeler has no idea why these results are being observed. Although computer programs must be as logically tight as the proof of a theorem, the program logic is processual rather than causal. Hence, we can know with certainty that a set of implications follow from a set of assumptions, without knowing why. Without explicit identification of the causal mechanisms, we cannot rule out the possibility that the results are nothing more than artifacts of particular modeling technologies or even bugs in the source code.

This is the dilemma of “double agents” posed by tremendous advances in the speed and power of inexpensive desktop computers. Computation allows formal models of methodological individualism to become much richer than the comparatively stylized mathematical models in game theory. However, this computational power has a double edge. ABC models can become so complicated that it is no longer possible to identify the causal mechanisms responsible for the results. Computational agents make it possible to identify the logical implications of a set of behavioral and structural assumptions that could not be mathematically derived. Yet these same agents can also obscure the underlying causal mechanisms by opening the door to rich theoretical elaboration.

The solution is the development of a standardized methodology of systematic theoretical elaboration based on step-by-step experimentation. The fundamental principle underlying this methodology is Einstein’s imperative to make models as simple as possible and no simpler than necessary. Modelers must resist the temptation to add bells and whistles that make the model more realistic or that satisfy an insatiable curiosity to find out what might happen if some new parameter were added. Nothing is gained by making a model more realistic if this also makes it impossible to identify the causal mechanisms underlying the correlations between inputs and outputs.

The second implication of this criticism is that computational results lack the generality of deductive conclusions (Holland 1995: 100). For this reason, computational models should not be used when a mathematical solution is available. However, mathematical models pay a high

price for the ability to generate deductive conclusions: they require simplifying assumptions in order to be mathematically tractable. While computational models risk becoming too complex to be understood, mathematical models risk becoming too simple to address important theoretical questions.

For example, Centola *et al.* (2005) show that universal enforcement of an unpopular norm can be an equilibrium, and that the structure of the network determines whether this equilibrium can be reached from a starting point that is far from equilibrium. It can be proven mathematically that if everyone enforces the obligation to enforce, then no one has an incentive to deviate unilaterally. However, this does not imply that this equilibrium will ever be reached. That question might also be answered mathematically for a randomly mixed or fully connected population. The problem is that few social networks are random and large populations are almost never fully connected. If we want to understand how an enforcement cascade might spread successfully on a complex network, then computational methods are likely to be required.

More generally, ABC models lack the generality of deductive conclusions, but they may nevertheless be the only practical method to capture the out-of-equilibrium dynamics that guide a population of agents from some initial condition to a stable social arrangement or from one equilibrium (when perturbed) to another. Agent-based models also extend equilibrium analysis to cases in which social stability arises even though individual strategies are constantly changing. A serious limitation of Nash equilibrium – the principal solution concept in analytical game theory – is the assumption that population stability is predicated upon the stability of the individual members, that is, the unwillingness of anyone to change strategies unilaterally. Young (1993, 1998) has proposed an alternative mathematical solution – stochastic stability – that overcomes this limitation of static equilibrium analysis. However, it remains to be shown that stochastic stability is mathematically tractable for games that are played in complex networks. Computational methods remain the most practical method for modeling dynamic equilibrium as well as out-of-equilibrium population behavior, especially for populations with complex network structures.

A dynamic equilibrium obtains when the forces pushing the system in one direction are precisely balanced by countervailing forces, as in models of system dynamics (Forrester 1968). ABC models, however, do not theorize these forces at the system level. Rather, the forces emerge out of the interactions of large numbers of interdependent but autonomous agents. For example, van de Rijdt *et al.* (2009) show that segregation can be an equilibrium in a population of unhappy multiculturalists. The

agents are moving at every opportunity, but the population level distribution remains homeostatic.

ABC models have another important benefit – the ability to model heterogeneity among the actors. It may be mathematically impossible to derive equilibria for a population of players with different preferences, different resources, and incomplete information about the preferences and resources of other players. Because ABC models are implemented at the individual level, they can be applied to large populations of heterogeneous agents just as easily as to populations in which everyone is exactly like everyone else and knows everything that everyone else knows.

In sum, mathematical models have an important advantage over ABC models in deriving conclusions deductively rather than by numerical computation. Whenever mathematically feasible, deductive methods should be preferred, e.g. in identifying Nash equilibria. However, deductive methods have limited application to out-of-equilibrium dynamics, especially for populations that are not randomly or fully connected. ABC models facilitate exploration of the effects of network topology on aggregate dynamics, such as the study of complex contagions by Centola and Macy (2007). Computation also invites rich elaboration of models that can make the models much more realistic than the highly stylized mathematical models used in game theory. This is not necessarily an advantage. Excessive complexity can obscure identification of the causal mechanisms that underlie correlations between parameter inputs and the behavior of the population. It may therefore be useful to use rich models as testbeds to discover regularities that can then be derived mathematically using a stripped down version of the model, as van de Rijt *et al.* did in their (2009) study of the segregative effects of linear and discontinuous ethnic preferences. Despite the inability to yield deductive generalizations, ABC models remain essential for the study of complex adaptive systems, whether social, physical or biological. As Holland concludes (1995: 195), “I do not think we will understand morphogenesis, or the emergence of organizations like Adam Smith’s pin factory, or the richness of interactions in a tropical forest, without the help of such models.”

REFERENCES

- Axelrod, Robert. 1997. *The Complexity of Cooperation*. Princeton University Press.
- Bruch, E. and R. Mare. 2006. “Neighborhood choice and neighborhood change,” *American Journal of Sociology* 112: 667–709.
2009. “Preferences and pathways to segregation: reply to van de Rijt, Siegel, and Macy,” *American Journal of Sociology* 114: 1181–98.

- Centola, D and M. Macy. 2005. "Social life in silico: the science of artificial societies," in Susan Wheelan (ed.), *Handbook of Group Research and Practice*. Thousand Oaks: Sage Publications, 273–83.
2007. "Complex contagions and the weakness of long ties," *American Journal of Sociology* 113: 702–34.
- Centola, D., R. Willer and M. Macy. 2005. "The emperor's dilemma: a computational model of self-enforcing norms," *American Journal of Sociology* 110: 1009–40.
- Flache, A. and M. Macy. 1996. "The weakness of strong ties: collective action failure in a highly cohesive group," *Journal of Mathematical Sociology* 21: 3–28.
2009. "Social dynamics from the bottom up: agent-based models of social interaction," in P. Bearman and P. Hedström (eds.), *Oxford Handbook of Analytical Sociology*. Oxford University Press.
- Flache, A., M. Macy and W. Raub. 2000. "Do company towns solve free rider problems? A sensitivity analysis of a rational-choice explanation," in W. Raub and J. Weesie (eds.), *The Management of Durable Relations: Theoretical and Empirical Models for Households and Organizations*. Amsterdam: Thela Thesis.
- Forrester, J. 1968. *Principles of Systems*. Waltham: Pegasus Communications.
- Gilbert, N. and K. Troitzsch. 1999. *Simulation for the Social Scientist*. Milton Keynes: Open University Press.
- Granovetter, M. 1973. "The strength of weak ties," *American Journal of Sociology* 78: 1360–80.
- Holland, J. 1995. *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Perseus.
- Kuran, T. 1995. *Private Truths, Public Lies: the Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Macy, M. and A. van de Rijt. 2006. "Game theory," in B. Turner (ed.), *The Cambridge Dictionary of Sociology*. Cambridge University Press.
- Macy, M. and R. Willer. 2002. "From factors to actors: computational sociology and agent-based modelling," *Annual Review of Sociology* 28: 143–66.
- Pancs, R. and N. Vriend. 2007. "Schelling's spatial proximity model of segregation revisited," *Journal of Public Economics* 91: 1–24.
- Prentice, D. and D. Miller. 1993. "Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm," *Journal of Personality and Social Psychology* 64: 243–56.
- Prietula, M., K. Carley and L. Gasser. 1998. *Simulating Organizations: Computational Models of Institutions and Groups*. Cambridge, MA: MIT Press.
- Schelling, T. 1971. "Dynamic models of segregation," *Journal of Mathematical Sociology* 1: 143–86.
- Simon, H. 1982. *Models of Bounded Rationality*. Cambridge, MA: MIT Press.
1998. *The Sciences of the Artificial*. Cambridge, MA: MIT Press.
- Van de Rijt, A., D. Siegel and M. Macy. 2009. "Neighborhood chance and neighborhood change: a comment on Bruch and Mare," *American Journal of Sociology* 114: 1166–80.

- Watts, D. and S. Strogatz. 1998. "Collective dynamics of 'small-world' networks," *Nature* 393: 409–10.
- Willer, R., K. Kuwabara and M. Macy. 2009. "The false enforcement of unpopular norms," *American Journal of Sociology* 115: 451–90.
- Wooldridge, M. and N. Jennings. 1995. "Intelligent agents: theory and practice," *Knowledge Engineering Review* 10: 115–52.
- Young, H.P. 1993. "The evolution of conventions," *Econometrica* 61: 57–84.
1998. *Individual Strategy and Social Structure: an Evolutionary Theory of Institutions*. Princeton University Press.
- Zhang, J. 2004. "Residential segregation in an all-integrationist world," *Journal of Economic Behavior and Organization* 54: 533–50.