

THE EMERGENCE OF COMPUTATIONAL SOCIOLOGY

NORMAN P. HUMMON* and THOMAS J. FARARO*

Department of Sociology, University of Pittsburgh, Pittsburgh, PA 15260, USA

The world of science has undergone a major transformation by virtue of technological innovations in computing and information processing. Sociology is one site in which this change is being played out. Our basic aim is to set out a revised image of any modern science, within which we can conceptualize and discuss the role of a newly emergent subfield we term computational sociology. Specifically, we expand the familiar two-component model of a science, featuring a theoretical and an empirical side, to include a computational component. We show how the three components interrelate in a triangular system in which empirical data analysis, theoretical explanation and computer simulation link the three components. We close our paper with a brief discussion of how one new development in computation relates to concepts of sociology, an instance of the hybrid character of computational sociology.

THE THREE COMPONENTS OF SCIENCE

Recent developments in the natural sciences are beginning to transform our image of the basic components of any science. The familiar two-component model of science as an interplay of theory and empirical research is being expanded to include a third component, computation. Physics, for instance, has theoretical physics and experimental physics, and now computational physics also is a social reality. Chemistry and biology have evolved and expanded in similar ways. With the new three-component model of any science (say, X), three types of expertise arise: theoretical X , empirical X , and computational X and roles develop for those who specialize in such activity.

Relations between the three components of a science exist and correspond to three modern scientific activities: explanation is the interplay of theoretical X and empirical X ; data analysis is the interplay of empirical X and computational X ; and simulation analysis is the interplay of theoretical X and computational X . Computational sociology, then has at least two roles, according to this model of science (Figure 1).

First, computational sociology uses the ideas and technologies of modern computer science to help advance theoretical sociology through the construction and study of simulation models. The simulation models are related to the world indirectly through the linkages with theoretical sociology and empirical sociology. Insofar as the fundamental ideas of the theories have gained acceptance in a scientific

*The authors acknowledge the extensive and useful editorial comments and suggestions provided by David Heise.

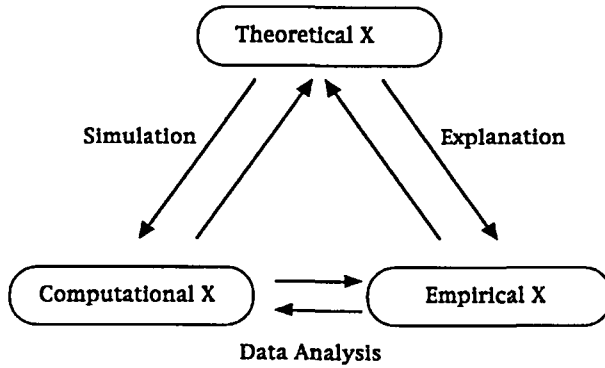


FIGURE 1. The three components of science.

community, to that extent the results of the simulation help to explain a body of empirical phenomena. Otherwise, what we learn from the simulation studies are essentially logical implications of our theoretical assumptions. These help further empirical research to test ideas and to advance our explanatory capabilities as a science.

The spread of personal computing and the acceleration of advances in both hardware and software have created a revival of interest in simulation. For basic research, it is the linkage to theory that makes simulation studies using the new technology so important. Among the growing number of contributions in this direction, an important example is the role of simulation studies in theory-driven research on power in social networks (Cook, et al., 1983; Marsden, 1983; Markovsky, 1987; Yamagishi, Gillmore and Cook, 1988; Skvoretz and Willer, 1991). Recognition of the general significance of simulation methodology for theoretical sociology is growing (see, for instance, Collins, 1988: Appendix).

Simulation is above all concerned with process. As Evans (1988: 19) defines the term, "a simulation is a dynamic model implemented on a computer." If the dynamic model is simple enough, then no simulation may be needed to study its properties. But the more usual scientific experience today is that simple rules have complex consequences when they involve concatenation over time. Thus, to implement a dynamic model on a computer is to create a computational model that mirrors the analytical model such that computational processes will produce the model outcomes for a variety of conditions. As we vary the conditions, the model outcomes vary in ways that we learn about through use of the simulation program. There are various types of simulation, such as those implementing continuous-change (Hanneman, 1988) and those implementing event-driven dynamics (Fararo and Hummon, 1994). Yet, in our terms, these are only variants on a common theme in which the emergent computational sociology functions as a handmaiden of theoretical sociology.

Procedures developed in the context of simulation have a further relevance through their indirect connection to empirical sociology. Once a simulation program is constructed, it constitutes software that enables computational processes in the

other key role of data analysis. Empirical researchers with certain forms of data can use the simulation programs to generate grounded measures and processes.

Thus, this brings us to the second role for computational sociology: using the ideas and technologies of modern computer science to help advance empirical sociology through the development of computational methods applicable to the reduction and analysis of empirical data. This empirical role for computational sociology has been evolving for more than three decades and includes at least five important capabilities: calculation; statistical simulation; data management; mathematical manipulation of data; and presentation of results.

First, sociologists were quick to recognize the importance of making statistical calculations by computer. The 1950 Census produced statistical tabulations by computer, and within a decade sociologists in major university research centers were conducting statistical calculations by computer.¹ Integrated statistical packages like BMD(P) and PSTAT, followed by DATATEXT, OSIRIS, SPSS and SAS, extended statistical computation to a much larger group sociologists in the 1960s and 1970s. In the last decade, increasingly powerful desktop computers and a new generation of software have taken over statistical computing for the discipline.

Second, like other statistically oriented scientists, sociologists are using statistical simulation to extend their empirical results beyond the limits of mathematical statistics. If we can't compute the properties of an estimator directly, we can simulate those properties using a variety of Monte Carlo based methods (Kennedy and Gentle, 1980).

Third, sociologists have adopted computer based data management technology. This has impacted empirical sociology in two ways. We can usefully employ far greater quantities of data, and thereby tackle data analyses beyond the scope of practicality even a few years ago. For example, desktop machines can analyze *all* annual surveys in the General Social Survey, from 1972 to the present; the GSS database is about 60 megabytes of data. Perhaps even more important than quantities of data is the ability to manage and use data stored in complex data structures. Modern relational database technology supports almost any data structure sociologists can conceive, e.g. complex kinship structures, or the hypergraph structures of contemporary network analysis (Date, 1987).

Fourth, some specialties of sociology make extensive use of computational mathematics. A prime example is network analysis, which applies graph theory and matrix algebra to the analysis of social structures. As in many other areas of science, these specialties of sociology would be impossible without computational mathematics (Burt, 1989; Borgatti, Everett, and Freeman, 1992).

Fifth, computer graphics is a new technology that is beginning to make an impact on sociology in two ways: the analysis of data; and the display of results. Statistical graphics techniques extend the capability of sociologists to analyze data in ways unknown during the early years of statistical computing. New desktop packages, "show" empirical relations in two, three, or higher order dimensions. These rela-

¹Personal communication with Phil Sidel of the University of Pittsburgh, who worked at the Bureau of the Census, Department of Commerce, and the Bureau of Applied Social Research at Columbia University during the 1950's.

a new programming development, object-orientation, and illustrated how such a purely technical aspect of computational sociology has suggestive correspondences with key concepts of contemporary sociology. This illustrates how computational sociology is an area in which advances in computation are merged with sociological ideas. The results, we hope, will exhibit the sort of hybrid vigor that can help sociology to accomplish its scientific goals.

REFERENCES

- Bainbridge, W. S., Brent, E. E., Carley, K. M., Heise, D. R., Macy, M. W., Markovsky, B. and Skvoretz, J. (1994) Artificial social intelligence. *Annual Review of Sociology* 10.
- Berger, J., Cohen, B. P., Snell, J. L. and Zelditch, M., Jr. (1962) *Types of Formalization*. Boston: Houghton-Mifflin.
- Borgatti, S., Everett, M. and Freeman, L. (1992) *UCINET IV, Version 1.0 Reference Manual*. Columbia, SC: Analytic Technologies.
- Burt, R. (1989) *Structure*. Center for Social Sciences, Columbia University, New York, NY.
- Cantu, M. and Tendon, S. (1992) *Borland C++ 3.1 Object-Oriented Programming*. New York: Bantam.
- Collins, R. (1988) *Theoretical Sociology*. San Diego: Harcourt Brace Jovanovich.
- Cook, K. S., Emerson, R. M., Gillmore, M. R. and Yamagishi, T. (1983) The distribution of power in exchange networks: Theory and experimental results. *American Journal of Sociology* 89: 275-305.
- Date, C. J. (1987) *Introduction to Database Systems*, vol. I, 4th ed. Reading, MA: Addison Wesley.
- Evans, J. B. (1988) *Structures of Discrete Event Simulation*. New York: Wiley.
- Fararo, T. J. and Hummon, N. P. (1994) Discrete event simulation and theoretical models in sociology, forthcoming in Markovsky, B., O'Brien, J. and Heimer, K. (eds.) *Advances in Group Processes*, vol. 11. Greenwich, CT: JAI Press.
- Fararo, T. J. and Skvoretz, J. (1986) E-state structuralism: A theoretical method. *American Sociological Review* 52: 591-602.
- Fararo, T. J. and Skvoretz, J. (forthcoming) Synthesizing theories of deviance and control with steps toward a dynamic sociocultural network model. *Journal of Mathematical Sociology* (special issue on Dynamic Social Networks, edited by F. Stokman and P. Doreian).
- Hanneman, R. A. (1988) *Computer-Assisted Theory Building: Modeling Dynamic Social Systems*, Newbury Park, CA: Sage.
- Heise, D. (1979) *Understanding Events*, New York: Cambridge University Press.
- Hummon, N. P. and Fararo, T. J. (1995) Actors and networks as objects. *Social Networks*, 17 1-26.
- Hummon, N. P. and Fararo, T. J. (this issue) Sociological algorithms for hierarchy and balance in dynamic network models.
- Kennedy, W. J., Jr. and Gentle, J. E. (1980) *Statistical Computing*, New York: Marcel Dekker Inc.
- MacKinnon, N. J. (1994) *Symbolic Interactionism as Affect Control*, Albany, NY: State University of New York Press.
- Markovsky, B. (1987) Toward multilevel sociological theories: Simulations of actor and network effects. *Sociological Theory* 5: 100-115.
- Marsden, P. V. (1983) Restricted access in networks and models of power. *American Journal of Sociology* 88: 686-717.
- Parsons, T. and Platt, G. M. (1973) *The American University*, Cambridge, MA: Harvard University Press.
- Powers, W. T. (1973) *Behavior: The Control of Perception*, Chicago: Aldine.
- Sedgewick, R. (1983) *Algorithms*, Addison Wesley, Reading, MA.
- Skvoretz, J. and Willer, D. (1991) Power in exchange networks: Setting and Structural variations. *Social Psychology Quarterly* 54: 224-238.
- Yamagishi, T., Gillmore, M. R. and Cook, K. S. (1988) Network connections and the distribution of power in exchange networks. *American Journal of Sociology* 93: 833-851.