Artificial Intelligence and Social Theory: A One-Way Street?

Sun-Ki Chai, University of Hawai'i

IN recent years, there has been a relatively small but increasingly influential set of research projects that import and adapt artificial intelligence models and techniques into sociology, most notably through the use of agent-based social simulations (see Bainbridge et al, 1994; Macy and Willer 2002 for overviews). This is a natural and welcome joining of two different disciplinary approaches to the modeling of human thought and behavior. However, for the most part, the flow of knowledge has unidirectional rather than reciprocal. While there have been a few notable papers that integrate mainstream social theory ideas with AI-style models (e.g. Carley 1991), there has been little attempt to export general social theories into artificial intelligence, either in the design of abstract software architectures or in the generation of practical working systems.

In many ways this is quite puzzling, since artificial intelligence has always been an area of research that has been centrally concerned with its theoretical foundations, which are meant to reflect deeper structures of human cognition and motivation. Furthermore, AI has readily adopted theoretical concepts and assumptions from other disciplines, including the social sciences. This dates from its earliest days, when it incorporated rules of inference and resolution from formal logic into the production rules of expert systems, and adopted structural and transformational grammars from linguistics not only predict the behaviors of their team members under realtime conditions of uncertainty.

In order for the environment to be useful, it is clear that the agents in the system must be endowed with cultural characteristics that closely mimic those of their human counterparts. At the same time, they must be modeled in such a way as to provide fairly decisive predictions about their behavior. Such systematic ways of representing culture do not appear in the artificial intelligence literature, hence they were adapted for this project from three different sources in social theory, taken from various social science disciplines. First of all, the representational framework for culture has been taken from the grid-group typology used in cultural

anthropology and political science (Douglas 1970, Thompson, et al. 1990). Propositions about cultural change have been adopted from my own "coherence" model, which in turn formalizes propositions from social psychological theories of attitude formation and social constructionist ideas from sociology (Chai 2001). Finally, the theory of action is drawn from rational choice theory, but is one in which the preferences and beliefs that drive the actions of individuals are not uniform, but rather seen as individual-level mappings of cultural differences. Without wholesale integration of ideas from social theories, it would frankly be difficult to envision a system with much predictive value.

Besides social simulation, another possible avenue for fruitful application of social theory to AI might be natural language processing, where an understanding of the social context of language is increasingly seen as key not only to unlocking semantics, but even syntax. Yet another is the design of user-modeling computer interfaces, since human-computer interaction are increasingly viewed as a form of social processes. In both of these areas, of course, a certain amount of formalization of sociological concepts will be necessary before they are ready for use in AI systems. On the other hand, one body of social theory that is already highly mathematical, social network theory, is more easily adaptable and indeed is already the subject of a number of applied projects that may in the future lead to working AI systems.

In one of these, 21st Century Systems and I are involved in work on intelligent software agents to automatically identify and analyze virtual communities on the web, in part through the use of network theory constructs.
into natural language processing systems, but also into the analysis of computing languages. Furthermore, throughout the years, there has been extensive sharing and collaboration between artificial intelligence and cognitive psychology communities on the theoretical foun-dations of perception and information processing.

Indeed, it is ironic that social simulation is perhaps the the area of artificial intelligence that is thinnest in its theoretical foundations. While it does utilize certain distinctive formalisms, most famously the cellular automata, assumptions about the bases for agent behavior within such simulations tend to be *ad hoc*. Agents are treated as extremely simple

entities following a few fixed strategies, which in turn are not based upon any underlying general model of human nature and society. For the most part, social simulations within sociology, rather than drawing on general social theory for their assumptions, have seemed to largely follow existing approaches from AI.

Admittedly, such approaches can reveal very interesting emergent properties, as simple strategies may lead to complex and counterintuitive outcomes at the macro level. Nonetheless, it is not unreasonable to raise the question of how outcomes might be different if agents were endowed with the arguably richer and more apparently human qualities found in micro-level social theories. Indeed, if we take the social theory enterprise seriously as a source for predicton, it follows that incorporating general social theory content into social simulations should allow artificial societies to better emulate real ones. By doing so, it should also likely render sociological work on simulation more relevant for the development of practical AI systems.

As an illustration of how such incorporation can occur, and without making any claims to definitiveness, I will discuss a project that I am working on in collaboration with the software company 21st Century Systems, Inc., under an Office of Naval Research grant. The objective of this project is to build a software module for the analysis of cultural differences. The module is designed for incorporation into a decision-support environment in which real world actors with whom the user is interacting are “avatarized” into agents whose movements appear within a graphical user interface. The purpose of the module is to help members of multinational coalitions operate better by helping them to

Whatever the realm, the potential uses of social theory to design intelligent software are numerous and remain largely untapped. For the immediate future, I would argue that artificial intelligence needs social theory as much or more than social theory needs artificial intelligence.

Bainbridge, William Sims, Edward E. Brent, Kathleen M. Carley, David R. Heise, Michael W. Macy, Barry Markovsky, and John Skvoretz. 1994. “Artificial Social Intelligence”, *Annual Review of Sociology* 20, 407-436.
Carley, Kathleen M. 1991. “A Theory of Group Stability”. *American Sociological Review* 56 (June), 331-54.
Chai, Sun-Ki. 2001. *Choosing an Identity: A General Model of Preference and Belief Formation*. Ann Arbor, MI: University of

Michigan Press.
Douglas, Mary. 1970. *Natural Symbols: Explorations in Cosmology*. New York: Pantheon.
Macy, Michael W. and Robert Willer. 2002. “From Factors To Actors: Computational Sociology and Agent-Based Modeling”. *Annual Review of Sociology* 28, 143-66.
Thompson, Michael, Richard Ellis and Aaron Wildavsky. 1990. *Cultural Theory*. Boulder: Westview Press.