

**COMPUTATIONAL MODELS OF COOPERATION: THE RELEVANCE OF
COGNITIVE PLAUSIBLE ACTOR MODELS FOR VALID MULTIAGENT
PERFORMANCE.**

Hans van den Broek
TNO Human Factors
PO Box 23
3769 ZG Soesterberg
The Netherlands
vandenbroek@tm.tno.nl

H.W.M Gazendam, J.L. Simons, and R.J.J.M. Jorna.

Faculty of Management and Organization
University of Groningen.
PO Box 800
9700 AV Groningen
The Netherlands

h.w.m.Gazendam@bdk.rug.nl; j.l.simons@bdk.rug.nl; r.j.j.m.Jorna@bdk.rug.nl

Topics: Docking of computational models; Computational modeling of social/organizational systems using cognitively realistic agents; Organizational adaptation and learning.

EXTENDED ABSTRACT

This paper presents a so-called alignment experiment of computational simulation models. An alignment study concerns the replication of a simulation experiment with the objective to establish whether two different simulation models claiming to deal with the same phenomena can produce the same results (Axtell, Axelrod, Epstein, & Cohen, 1995). Axelrod (1997) provides three reasons why replication of simulation experiments is important. (a) It is needed to confirm whether the claimed results of a given simulation are reliable in the sense that someone starting from scratch can reproduce them. Without this confirmation, it is possible that some published results are simply mistaken due to programming errors, misinterpretation of what was actually simulated, or errors in analyzing or reporting the results. (b) Replication can also be useful for testing the robustness of inferences from models. Finally, (c) replication is needed to determine if some model can subsume another. The incentive for alignment study matches the third category. The aim of the alignment study was to compare the performance of a multiagent simulation model using cognitive plausible agents with one based on cognitive implausible agents.

The simulation experiment that has been replicated within the alignment study has been described in “*When can two heads learn better than one? Results from a computer model of organizational learning*” (Cohen, 1992). The computer simulation experiment conducted by Cohen consists of (i) an actor model, capable of interacting with (ii) a task environment in order to carry out (iii) a task. The simulation model was set up by Cohen in order to answer the question whether a group can do tasks better than are, in principle, within the competence of a single agent. Hence, the basic experimental

design of the original experiment consists of contrasting the task performance of a single agent with the task performance of an organization consisting of two agents. The simulation results described by Cohen clearly indicate that the learning of multiple agents outperforms single agent learning. Hence, the Cohen experiment demonstrated that a task that in principle can be accomplished by a single agent was outperformed in collaborative effort. The learning engine used in the original experiment concerns a low-level learning system. It consists of matrix updating mechanism that summarizes previous experiences in that the entries of the matrix reflect the earlier reinforcement of associations between features and actions. Whenever an action moves the system to a new state that is closer to its goal (the agent receives explicit feedback about this) the association of that action with the features that were present in the prior state is strengthened; otherwise, it is weakened. The action taken is the one that has the strongest total of associations across all the features that the agent detects in the current state.

The replicated experiment mimics the original as close as possible except for the cognitive agents. Cognitive plausibility of the agents has been achieved by applying the Soar architecture (Laird, Newell, & Rosenbloom, 1987). The assumption is that Soar provides a computational and theoretical framework that enables to implement a cognitive plausible agent because Soar implements the unified theories of cognition as expressed by Newell (1990).

The alignment study has been conducted in order to validate the thesis that multiagent system validity, i.e. the external validity of the multiagent behavior, is based on agent validity (Van den Broek, in press). Multiagent simulation tools for organizations fit the theoretical framework which focuses on research of the empirical relations among (1) self-interested actions of individual actors; (2) the structural and organizational conditions coordinating these actions (voting procedures, markets, hierarchies, coordination of mutual understanding); and (3) the macro social effect resulting from these conditions (Carley & Prietula, 1994). Multiagent simulation models are coordinated problem solving models, which provide plausible simulation tools for addressing this theoretical framework because they relate the properties of the agents, the laws of their interaction, to the properties of the multiagent system as a whole.

On the multiagent level, we found remarkable differences between the two experimental versions. The Vandebroek version of the experiment incorporates the possibility of different interpretations of reality within the multiple agent decision-making model (Van den Broek & Gazendam, 1997). Because of that, “conflicts” in the collaborative effort emerged in the form of action-repetition deadlocks. It means that *both* agents persist in applying different actions to a particular problem solving state, which result in a cyclic action pattern. The disagreements occur because the model of the environment both agents learn to create may differ concerning their representational contents at some point in time, because the models of the task environment are based on individual experience of action-feedback cycles. Hence, inconsistencies of behavior emerge because of local differences of perception of reality. Each time cooperative conflicts emerge, the agents must put in effort in order to solve the conflict. The effort concerns the detection of the problem and by applying a coordination rule, which requires applying additional actions upon the task environment. The total effort needed for solving cooperative conflicts is viewed as coordination costs. Because cooperation

costs are involved, the cooperative work condition of the replicated experiment performs less well than the non-cooperative condition.

In contrast, within the Cohen experiment neither conflict nor coordination cost were involved within the collaborative effort. The conclusion therefore is that with incorporating the possibility of different interpretations of reality within the multiple agent decision-making model, the plausibility of the model enhances. Differences in perception of reality are identified as one of the causes of individual conflict (March & Simon, 1993: 141).

On the theoretical level, the conclusion of the alignment study is that because multiactor systems are interactional models, which explicitly link the information processing of the individual with the emerging process of coordination, one needs cognitive, i.e. representational agents in order to connect with coordination processes. Coordination processes are social processes and, therefore, agents need to be able to model the other agents within the task environment in addition of modeling the task environment itself. This means in terms of Carley and Newell (1994) that the agents need to be social agents, i.e. the agents need to be able to interact with others in order to bring about coordination.

LITERATURE:

- Carley, K. M., & Newell, A. (1994). The Nature of the Social Agent. *Journal of mathematical Sociology*, 19(4), 221-262.
- Carley, K. M., & Prietula, M. J. (1994). ACTS theory: extending the model of bounded rationality. In K. M. Carley & M. J. Prietula (Eds.), *Computational Organization Theory*. Hillsdale, NJ. Hove, UK.: Lawrence Erlbaum.
- Laird, J., Newell, A., & Rosenbloom, P. (1987). SOAR: an architecture for general intelligence. *Artificial Intelligence*, 33, 1-64.
- March, J. S., & Simon, H. A. (1993). *Organizations*. (Second ed.). Cambridge, Massachusetts: Blackwell Business.
- Newell, A. (1990). *Unified theories of cognition (The William James lectures; 1987)*. Cambridge, Massachusetts. London, England: Harvard University Press.
- Van den Broek, J. (in press). *Multiactor Simulation Models of Organizations: The Relevance of Cognitive Plausible Actor Models for Valid Multiactor Performance*. , University of Groningen, Groningen.
- Van den Broek, J., & Gazendam, H. W. M. (1997). Organizational actors and the need for a flexible world representation. In R. Conte, R. Hegselmann, & P. Terna (Eds.), *Simulating social phenomena* (Vol. 456, pp. 303-311). Berlin etc.: Springer.