

# A Cognitively Based Simulation of Simple Organizations

Ron Sun (rsun@rpi.edu)  
Isaac Naveh (yizchaknaveh@yahoo.com)  
Cognitive Science Department  
Rensselaer Polytechnic Institute  
Troy, NY 12180, USA

## Abstract

This paper explores cognitively realistic social simulations by deploying the CLARION cognitive architecture in a simple organizational simulation, which involves the interaction of multiple cognitive agents. It argues for an integration of the two separate strands of research: cognitive modeling and social simulation. Such an integration could, on the one hand, enhance the accuracy of social simulation models by taking into full account the effects of individual cognitive factors, and on the other hand, it could lead to greater explanatory, predictive, and prescriptive power from these models.

**Keywords:** social simulation; organization; decision making; cognitive architecture.

## Introduction

Most of the current work in social simulation still assumes very rudimentary cognition on the part of the agents (e.g., Gilbert and Doran 1994). At the same time, while researchers in cognitive science have devoted considerable attention to the workings of individual cognition (e.g., Anderson and Lebiere 1998; Sun 2002), socio-cultural processes and their relations to individual cognition have generally not been sufficiently studied by cognitive scientists (with some notable exceptions).

However, there are reasons to believe that better models of individual cognition can lead us to a better understanding of aggregate processes involving multi-agent interaction (Moss 1999; Castelfranchi 2001; Sun 2001). Cognitive models that incorporate realistic tendencies, biases, and capacities of individual cognitive agents can serve as a more realistic basis for understanding multi-agent interaction. Social interaction is, after all, the result of individual cognition (which includes instincts, routines, and patterned behavior, as well as complex symbolic, conceptual processes). Therefore, the mechanisms underlying individual cognition cannot be ignored in studying multi-agent interactions. At least, the implications of these mechanisms should be understood before they are abstracted away.

In the remainder of this paper, first, an existing organizational decision task is introduced. Then, a more realistic cognitive architecture, named CLARION, is described, which is applied to the organizational task. The idea is to substitute the more sophisticated agent, based on CLARION, for the simple agents used in the previous work, with the aim of studying the interaction of cognition and organizational design.

## Organizational Design

Carley et al (1998) introduced an organizational decision task involving different types of organizational structures and agents. By varying agent type and structure separately, they studied how these factors interact with each other.

**The Task.** The task is to determine whether a blip on a screen is a hostile aircraft, a flock of geese, or a civilian aircraft (Carley et al 1998). It has been extensively used before in studying organizational design.

In each case, there is a single object in the airspace. The object has nine different attributes, each of which can take on one of three possible values (e.g., its speed can be low, medium, or high). An organization must determine the status of an observed object: whether it is friendly, neutral or hostile. There are a total of 19,683 possible objects, and 100 problems are chosen randomly (without replacement) from this set. The true status of an object is determinable by adding up all nine attribute values. If the sum is less than 17, then it is friendly; if the sum is greater than 19, it is hostile; otherwise, it is neutral.

No one single agent has access to all the information necessary to make a choice. Decisions are made by integrating separate decisions made by different agents, each of which is based on a different subset of information.

In terms of organizational structures, there are two archetypal structures of interest: (1) teams, in which agents act autonomously, individual decisions are treated as votes, and the organization decision is the majority decision; and (2) hierarchies, which are characterized by agents organized in a chain of command, such that information is passed from subordinates to superiors, and the decision of a superior is based solely on the recommendations of his/her subordinates. In this task, only a two-level hierarchy with nine subordinates and one supervisor is considered.

In addition, organizations are distinguished by the structure of information accessible by each agent. There are two varieties of information access: (1) distributed access, in which each agent sees a different subset of three attributes (no two agents see the same set of three attributes), and (2) blocked access, in which three agents see exactly the same attributes. In both cases, each attribute is accessible to three agents.

Several simulation models were considered in Carley et al (1998). Among them, CORP-ELM produced the most

probable classification based on an agent’s own experience, CORP-P-ELM stochastically produced a classification in accordance with the estimate of the probability of each classification based on the agent’s own experience, CORP-SOP followed organizationally prescribed standard operating procedure (which involved summing up the values of the attributes available to an agent) and thus was not adaptive, and Radar-Soar was a (somewhat) cognitive model built in Soar, which is based on explicit, elaborate search in problem spaces (Rosenbloom et al 1991).

**Previous Experimental Results.** The experiments by Carley et al (1998) were done in a 2 x 2 fashion (organization x information access). In addition, human data for the experiment were compared to the results of the four aforementioned models (Carley et al 1998). The data appeared to show that agent type interacted with organizational design. See Table 1.

Agent/Org.	Team(B)	Team(D)	Hierarchy(B)	Hierarchy(D)
Human	50.0	56.7	46.7	55.0
Radar-Soar	73.3	63.3	63.3	53.3
CORP-P-ELM	78.3	71.7	40.0	36.7
CORP-ELM	88.3	85.0	45.0	50.0
CORP-SOP	81.7	85.0	81.7	85.0

Table 1: Human and simulation data for the organizational decision task. D indicates distributed information access, while B indicates blocked information access. All numbers are percent correct.

The human data showed that humans generally performed better in team situations, especially when distributed information access was in place. Moreover, distributed information access was generally better than blocked information access.

It also suggested that which type of organizational design exhibit the highest performance depends on the type of agent. For example, human subjects performed best as a team with distributed information access, while Radar-Soar and CORP-ELM performed the best in a team with blocked information access.

In their work, the agent models used were very simple. The “intelligence” level in these models was rather low (including the Soar model, which essentially encoded a set of simple rules). Moreover, learning in these simulations was rudimentary: there was no complex learning process as one might observe in humans.

## The Model

Below, we discuss a more cognitively realistic model, to be used for addressing this task (Sun 2002). First, some major issues that the model captures are as follows.

**Explicit vs. Implicit Learning.** The role of implicit learning in skill acquisition has been widely recognized in recent years (e.g., Reber 1989; Seger 1994; Stadler and Frensch 1998). Although explicit and implicit learning have both been actively studied, the question of the interaction between these two processes has rarely been broached. However, despite the lack of study of this interaction, it has recently become evident (e.g., in Seger 1994) that rarely, if ever, is only one of type of

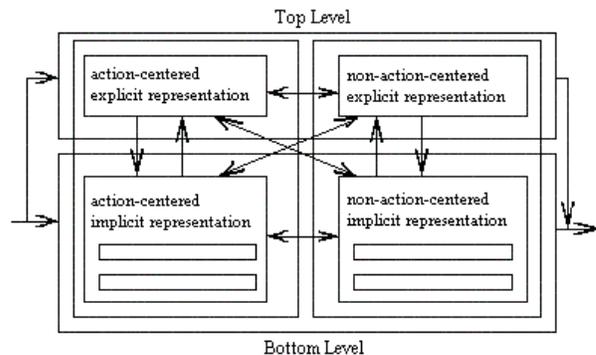


Figure 1: The CLARION architecture

learning engaged. Our review of experimental data (e.g., Reber 1989; Stanley et al 1989; Sun et al 2001) shows that although one can manipulate conditions so that one or the other type of learning is emphasized, both types of learning are nonetheless usually present.

To model the interaction between these two types of learning, the cognitive architecture CLARION was developed (Sun et al 2001), which captures the combination of explicit and implicit learning. CLARION mostly learns in a “bottom-up” fashion, by extracting explicit knowledge from implicit knowledge (see Sun 2002 for details). Such processes have also been observed in humans (e.g., Stanley et al 1989; Mandler 1992).

**A Sketch of the CLARION Model.** CLARION is an integrative cognitive architecture with a dual representational structure (Sun et al 1998; Sun et al 2001; Sun 2002). It consists of two levels: a top level that captures explicit learning, and a bottom level that captures implicit learning (see Figure 1).

At the bottom level, the inaccessibility of implicit learning is captured by subsymbolic distributed representations. This is because representational units in a distributed environment are capable of performing tasks but are generally not individually meaningful (Sun 1995). Learning at the bottom level proceeds in trial-and-error fashion, guided by reinforcement learning (i.e., Q-learning) implemented in backpropagation neural networks (Sun 2002).

At the top level, explicit learning is captured by a symbolic representation, in which each element is discrete and has a clearer meaning. This accords well with the directly accessible nature of explicit knowledge (Sun 2002). Learning at the top level proceeds by first constructing a rule that corresponds to a “good” decision made by the bottom level, and then refining it (by generalizing or specializing it), mainly through the use of an “information gain” measure that compares the success ratio of various modifications of the current rule.

A high-level pseudo-code algorithm that describes the action-centered subsystem of CLARION is as follows:

1. Observe the current state  $x$ .
2. Compute in the bottom level the Q-value of each of the possible actions ( $a_i$ 's) associated with the state  $x$ :  $Q(x, a_1)$ ,

$Q(x, a_2), \dots, Q(x, a_n)$ .

3. Find out all the possible actions ( $b_1, b_2, \dots, b_m$ ) at the top level, based on the state  $x$  and the rules in place at the top level.
4. Compare the values of  $a_i$ 's with those of  $b_j$ 's, and choose an appropriate action  $a$ .
5. Perform the action  $a$ , and observe the next state  $y$  and (possibly) the reinforcement  $r$ .
6. Update the bottom level in accordance with the *Q-Learning-Backpropagation* algorithm, based on the feedback information.
7. Update the top level using the *Rule-Extraction-Refinement* algorithm.
8. Go back to Step 1.

At the bottom level, a Q-value is an evaluation of the "quality" of an action in a given state:  $Q(x, a)$  indicates how desirable action  $a$  is in state  $x$ . Actions can be selected based on Q-values. To acquire the Q-values, Q-learning, a reinforcement learning algorithm (Watkins 1989), is used. (See Sun 2002 for further details.)

In the top level, explicit knowledge is captured in a simple prepositional rule form. We devised an algorithm for learning explicit knowledge (rules) using information from the bottom level (the *Rule-Extraction-Refinement*, or RER, algorithm). The basic idea is as follows: if an action decided by the bottom level is successful then the agent extracts a rule (with its action corresponding to that selected by the bottom level and with its conditions corresponding to the current state), and adds the rule to the top level. Then, in subsequent interactions with the world, the agent refines the extracted rule by considering the outcome of applying the rule: if the outcome is successful, the agent may try to generalize the conditions of the rule to make it more universal. If the outcome is unsuccessful, the agent may try to specialize the rule, by narrowing its conditions down and making them exclusive of the current state.

The information gain (IG) measure of a rule is computed (in this organizational decision task) based on the immediate feedback at every step when the rule is applied. The inequality,  $r > threshold_{RER}$  determines the positivity/negativity of a step and the rule matching this step (where  $r$  is the feedback received by an agent). The positivity threshold (denoted  $threshold_{RER}$  above) corresponds to whether or not an action is perceived by the agent as being reasonably good. Based on the positivity of a step, PM (Positive Match) and NM (negative match) counts of the matching rules are updated. IG is calculated based on PM and NM:

$$IG(A, B) = \log_2 \frac{PM_a(A) + c1}{PM_a(A) + NM_a(A) + c2} - \log_2 \frac{PM_a(B) + c1}{PM_a(B) + NM_a(B) + c2}$$

where A and B are two different rule conditions that lead to the same action  $a$ , and  $c1$  and  $c2$  are two constants representing the prior (by default,  $c1 = 1$ ,  $c2 = 2$ ). Essentially, the measure compares the percentages of positive matches under alternative conditions A and B.

The generalization operator is based on the IG measure. Generalization amounts to adding an additional value to one input dimension in the condition of a rule, so that the rule will have more opportunities of matching input. For a rule to be generalized, the following must hold:

$$IG(C, all) > threshold_{GEN} \quad \text{and} \quad max_{C'} IG(C', C) \geq 0$$

where  $C$  is the current condition of a rule (matching the current state and action), *all* refers to the corresponding match-all rule (with the same action as specified by the original rule but an input condition that matches any state), and  $C'$  is a modified condition equal to  $C$  plus one input value. If the above holds, the new rule will have the condition  $C'$  with the highest IG measure. The generalization threshold (denoted  $threshold_{GEN}$  above) determines how readily an agent will generalize a rule.

The specialization operator works in an analogous fashion, except that a value in an input dimension is discarded, rather than being added. In addition, to avoid the proliferation of useless rules, a RER *density* measure is in place. A density of  $1/x$  means that a rule must be invoked once per  $x$  steps to avoid deletion due to disuse. This corresponds to the agent's memory for rules, necessitating that a rule come up every once in a while in order to be retained.

To integrate results, levels are chosen stochastically, using a probability of selecting each level.

When the outcome from the bottom level is chosen, a stochastic process based on the Boltzmann distribution of Q values is used for selecting an action:

$$p(a|x) = \frac{e^{Q(x,a)/t}}{\sum_i e^{Q(x,a_i)/t}}$$

where  $x$  is the current state,  $a$  is an action, and  $t$  controls the degree of randomness (temperature) of the process.<sup>1</sup>

Below, we present three simulations involving the CLARION model. The first experiment uses the aforementioned radar task (Carley et al 1998) but substitutes a different cognitive model. In CLARION, we can easily vary parameters and options that correspond to different cognitive capacities and test the resulting performance. The second simulation uses the same task, but extends the duration of training given to the agents. Finally, in the third simulation, we vary a wide range of cognitive parameters of the model in a factorial design.

## Simulation I: Matching Human Data

In this simulation, each agent (whatever its position in the organization) is implemented as a CLARION model. At the top level, RER is used to extract rules. At the bottom level, each agent has a single Q-learning neural network that is trained, over time, to respond correctly. The network receives an external feedback of 0 or 1 after each step, depending on whether the target was correctly

<sup>1</sup>This method is also known as Luce's choice axiom (Watkins 1989). It is found to match psychological data in many domains.

classified by the network. Various parameter values were chosen through trial-and-error optimization.

The results of our simulation are shown in Table 2. 3,000 training cycles (each corresponding to a single instance, followed by a decision by the organization) were used for each organization in this simulation. The agents of an organization were trained together within that organization. Other settings (such as organizational structures and information access) were the same as in Carley et al (1998) as described earlier. As can be seen, our results closely accord with the patterns of the human data, with teams outperforming hierarchal structures, and distributed access proving superior to blocked access. Also, as in humans, performance is not grossly skewed towards one condition or the other, but is roughly comparable across all conditions (unlike some of the simulation results from Carley et al 1998). The match with the human data is far better than in the simulations conducted in the original study (Carley et al 1998). The better match is due, at least in part, to a higher degree of cognitive realism in our simulation.

Agent/Org.	Team(B)	Team(D)	Hierarchy(B)	Hierarchy(D)
Human	50.0	56.7	46.7	55.0
CLARION	53.2	59.3	45.0	49.4

Table 2: Simulation data for agents running for 3,000 cycles. The human data from Carley et al (1998) are reproduced here. Performance of CLARION is computed as percent correct over the last 1,000 cycles.

Because no further details of the human data and the models of Carley et al (1998) are available, no further comparisons to the CLARION simulation can be made.

## Simulation II: Extending the Previous Simulation

So far, we have considered agents trained for only 3,000 cycles. The results were interesting, because they were analogous to those of humans. However, it is interesting to see what will happen if we extend the length of the training. In particular, we are interested in knowing if the trends seen above will be preserved in the long run. It is important that before we draw any conclusion about human performance, we understand the context and conditions under which data are obtained, and thereby avoid over-generalizing possible conclusions (e.g., team vs. hierarchy, blocked vs. distributed; Carley et al 1998).

Figures 2-5 show learning as it occurs over 20,000 (rather than 3,000) cycles. Previously, the best-performing condition was team organization with distributed information access. As can be seen in Figure 2, this condition continues to improve slowly after the first 3,000 cycles. However, it is overtaken by team organization with blocked access (Figure 3). Thus, it seems that while teams benefit from a diversified (distributed) knowledge base in the initial phase of learning, a well-trained team with redundant (blocked) knowledge performs better in the long run.

In the hierarchal conditions, too, we can see either a reversal or disappearance of the initial trends. Hierar-

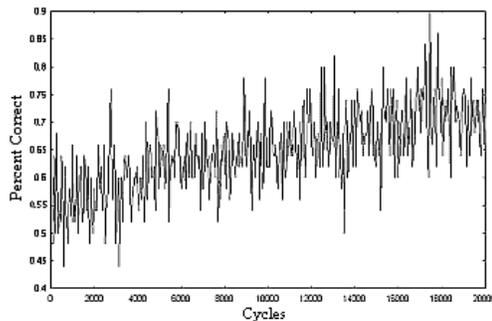


Figure 2: Training curve for team organization with distributed access

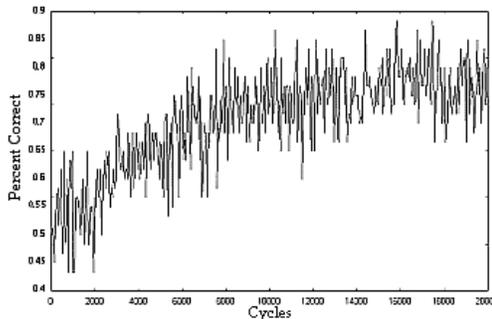


Figure 3: Training curve for team organization with blocked access

chies using distributed access (Figure 4) now show not only the best, but also the most stable (least variance) performance of any condition. Likewise, a hierarchy with blocked access (Figure 5), previously a weak performer, shows impressive gains in the long run. Thus, while hierarchies take longer to train, their performance is superior in the long run. In a hierarchy, a well-trained supervisor is capable of synthesizing multiple data points with greater sensitivity than a simple voting process. Likewise, the reduced individual variation in blocked access leads to less fluctuation in performance in the long run.

There is a serious lesson here: limited data can allow us to draw only limited conclusions—only with regard to the specific situation under which the data were obtained. There is a natural tendency for researchers to over-generalize their conclusions (Carley et al 1998), which can only be remedied by more extensive investigations. Given the high cost of human experiments, simulation has a large role to play in exploring alternatives and possibilities, especially social simulation with cognitive architectures.

## Simulation III: Varying Cognitive Parameters

In the two preceding simulations, agents were run under a fixed set of cognitive parameters. Next let us see what happens when we vary these parameters (analogous to

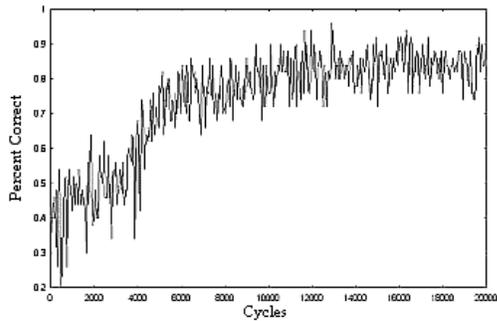


Figure 4: Training curve for hierarchal organization with distributed access

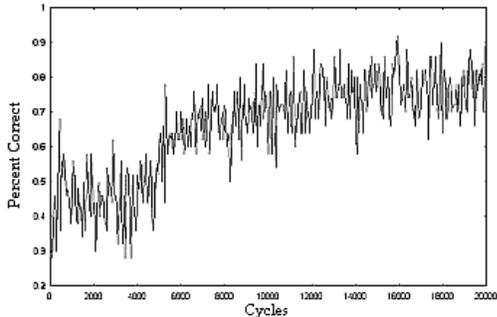


Figure 5: Training curve for hierarchal organization with blocked access

varying the training length earlier). This again allows us to see the variability of results, and thus avoid over-generalization. Because CLARION captures a wide range of cognitive processes and phenomena, its parameters are generic rather than task-specific. Thus, we have the opportunity of studying specific issues (such as organizational design), in the context of a general theory of cognition.

Two sets of parameters of CLARION were separately varied (in order to avoid the prohibitively high cost of varying all parameters simultaneously). The first set of parameters consisted of fundamental parameters of CLARION: (1) Reliance on the top versus the bottom level, expressed as a fixed probability of selecting each level. (2) Learning rate of the neural networks. (3) Temperature, or degree of randomness. The second set consisted of parameters related to RER, including: (1) RER positivity threshold, which must be exceeded for a rule to be considered “successful.” (2) RER density measure, which determined how often a rule must be invoked in order to be retained. (3) RER generalization threshold, which must be exceeded for a rule to be generalized. (These parameters were also described earlier.)

An ANOVA on the results confirmed the effects of organization [ $F(1, 24) = 30.28, p < 0.001, MSE = 0.05$ ] and information access [ $F(1, 24) = 7.14, p < 0.05, MSE = 0.01$ ] to be significant. Moreover, the interaction of these two factors with length of training was significant [ $F(1,$

$24) = 59.90, p < 0.001, MSE = 0.73$  for organization;  $F(1, 24) = 3.43, p < 0.05, MSE = 0.01$  for information access]. These interactions reflect the trends discussed above: the superiority of teams and distributed information access at the start of the learning process, and either the disappearance or reversal of these trends towards the end. This finding is important, because it shows that these trends persist robustly across a wide variety of settings of cognitive parameters, and do not critically depend on any one setting of these parameters.

The effect of probability of using the top vs. the bottom level was likewise significant [ $F(2, 24) = 11.73, p < 0.001, MSE = 0.02$ ]. More interestingly, however, its interaction with length of training was significant as well [ $F(2, 24) = 12.37, p < 0.001, MSE = 0.01$ ]. Rule learning is far more useful at the early stages of learning, when increased reliance on them tends to boost performance, than towards the end of the learning process. This is because rules are crisp guidelines that are based on past success, and as such, they provide a useful anchor at the uncertain early stages of learning. However, by the end of the learning process, they become no more reliable than highly-trained neural networks. This corresponds to findings in human cognition, where there are indications that rule-based learning is widely used in the early stages of learning, but is later increasingly supplanted by similarity-based processes (Smith and Minda 1998) and skilled performance (Anderson and Lebiere 1998). Such trends may partially explain why hierarchies do not perform well initially because a hierarchy’s supervisor is burdened with a higher input dimensionality, and therefore it takes a longer time to encode rules (which are essential at the early stages of learning).

Predictably, the effect of learning rate was significant [ $F(2, 24) = 32.47, p < 0.001, MSE = 0.07$ ]. Groups with a higher learning rate (0.5) outperformed the groups with the lower learning rate (0.25) by between 5-14%. However, there was no significant interaction between learning rate and organization or information access. This suggests that quicker learners do not differentially benefit from, say, a hierarchy versus a team. By the same token, the poorer performance of slower learners cannot be mitigated by recourse to a particular combination of organization and information access.

Let us now turn to the parameters related to RER rule learning. Generalization threshold determines how readily an agent will generalize a successful rule. It is better to have a higher rule generalization threshold than a lower one (up to a point). An ANOVA confirmed the significance of this effect [ $F(1, 24) = 15.91, p < 0.001, MSE = 0.01$ ]. Thus, if one restricts the generalization of rules only to those rules that have proven relatively successful (by selecting a fairly high generalization threshold), the result is a higher-quality rule set, which leads to better performance in the long run.

Relatedly, while the effect of rule density on performance was insignificant, the interaction between density (i.e., “memory” for rules) and generalization threshold was significant [by an ANOVA;  $F(2, 24) = 2.93; p < 0.05; MSE = 0.01$ ]. When rules are of relatively high

quality (i.e., under a high generalization threshold) it is advisable to have more of them available (which is achievable by lowering the density parameter). By contrast, when the average quality of rules is lower (i.e., under a low generalization threshold) it is advantageous to have a quicker forgetting process in place, as embodied by a high density parameter.

Finally, the interaction between generalization threshold and organization was significant at the start of the learning process [by an ANOVA;  $F(1, 24) = 5.93$ ,  $p < 0.05$ ,  $MSE = 0.01$ ], but not at the end. This result is more difficult to interpret, but probably reflects the fact that hierarchies, at the start of the learning process, do not encode very good rules to begin with (due to the higher input dimensionality of the supervisor and the resulting learning difficulty). Thus, generalizing these rules, even incorrectly, causes relatively little further harm.

This simulation confirmed an earlier observation—namely, that which organizational structure (team vs. hierarchy) or information access scheme (distributed vs. blocked) is superior depends on the length of the training. It also showed that some cognitive parameters (e.g., learning rate) have a monolithic, across-the-board effect under all conditions, whereas in other cases, complex interactions of factors are at work. See Sun and Naveh (2004) for full details. This illustrates, once again, the importance of limiting one's conclusions to the specific cognitive context in which data were obtained.

## Discussion

By using CLARION, we have been able to more accurately capture organizational performance data as well as to formulate deeper explanations for the results observed, due to cognitive realism. For instance, based on our observations, one may formulate the following possible explanation: the poorer performance of hierarchies early on (see Simulation I) may be due, at least in part, to the longer training time needed to encode high-dimensional information by the supervisor, which leads to fewer useful rules being acquired at the top level. This in turn impacts performance, since rule learning is especially important in the early stages of learning (see Simulation III). Such explanations are only possible when the model is cognitively realistic. Beside offering deeper explanations, cognitive realism may also lead to greater predictive and prescriptive power for social simulations. In CLARION, we can vary parameters and options that correspond to cognitive processes and test their effect on performance. In this way, CLARION can be used to predict human performance in organizations, and furthermore to help performance by prescribing optimal or near-optimal cognitive abilities for specific tasks and organizational structures (see Sun and Naveh 2004).

## References

Anderson, J. R. and Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Carley, K. M., Prietula, M. J., and Lin, Z. (1998). Design versus Cognition: The interaction of agent cognition and organizational design on organizational performance. *Journal of Artificial Societies and Social Simulation*, 1(3).
- Castelfranchi, C. (2001). "The Theory of Social Functions: Challenges for Computational Social Science and Multi-Agent Learning." In Sun R (ed.), *Cognitive Systems Research*, special issue on the multidisciplinary studies of multi-agent learning, 2(1). pp. 5-38.
- Gilbert, N. and Doran, J. (1994). *Simulating Societies: The Computer Simulation of Social Phenomena*. London, UK: UCL Press, London.
- Mandler, J. (1992). How to Build a Baby. *Psychological Review*, 99(4), pp. 587-604.
- Moss, S. (1999). Relevance, Realism and Rigour: A Third Way for Social and Economic Research. *CPM Report No. 99-56*. Manchester, UK: Center for Policy Analysis, Manchester Metropolitan University.
- Reber, A. (1989). Implicit Learning and Tacit Knowledge. *Journal of Experimental Psychology: General*, 118(3). pp. 219-235.
- Rosenbloom, P., Laird, J., Newell, A., and McCarl, R. (1991). A Preliminary Analysis of the Soar Architecture as a Basis for General Intelligence. *Artificial Intelligence*, 47(1-3). pp. 289-325.
- Seger, C. (1994). Implicit Learning. *Psychological Bulletin*, 115(2). pp. 163-196.
- Smith, J. D. and Minda, J. P. (1998). Prototypes in the Mist: The Early Epochs of Category Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24. pp. 1411-1436.
- Stadler, M. and Frensch, P. (1998). *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage Publications.
- Stanley, W., Mathews, R., Buss, R., and Kotler-Cope, S. (1989). Insight Without Awareness: On the Interaction of Verbalization, Instruction and Practice in a Simulated Process Control Task. *Quarterly Journal of Experimental Psychology*, 41A(3). pp. 553-577.
- Sun, R. (2001). Cognitive Science Meets Multi-Agent Systems: A Prolegomenon. *Philosophical Psychology*, 14(1). pp. 5-28.
- Sun, R. (2002). *Duality of the Mind*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Sun, R., Merrill, E., and Peterson, T. (2001). From Implicit Skills to Explicit Knowledge: A Bottom-Up Model of Skill Learning. *Cognitive Science*, 25(2). pp. 203-244.
- Sun, R. and Naveh, I. (2004). Simulating Organizational Decision-Making Using a Cognitively Realistic Agent Model. *Journal of Artificial Societies and Social Simulation*, Vol.7, No.3.
- Watkins, C. (1989). *Learning with Delayed Rewards*. PhD Thesis, Cambridge University, Cambridge, UK.