A Cognitively Based Simulation of Academic Science

Isaac Naveh
Department of Computer Science
University of Missouri
Columbia, MO 65211, USA

Ron Sun
Department of Cognitive Science
Rensselaer Polytechnic Institute
Troy, NY 12180, USA
Phone: (518) 276 3409
Fax: (518) 276 8268
Email: rsun@rpi.edu

Abstract

The models used in social simulation to date have mostly been very simplistic cognitively, with little attention paid to the details of individual cognition. This work proposes a more cognitively realistic approach to social simulation. It begins with a model created by Gilbert (1997) for capturing the growth of academic science. Gilbert's model, which was equation-based, is replaced here by an agent-based model, with the cognitive architecture CLARION providing greater cognitive realism. Using this cognitive agent model, results comparable to previous simulations and to human data are obtained. It is found that while different cognitive settings may affect the aggregate number of scientific articles produced, they do not generally lead to different *distributions* of number of articles per author. The paper concludes with a discussion of the correspondence between our model and the constructivist view of academic science. It is argued that using more cognitively realistic models in simulations may lead to novel insights.

1 Introduction

Computer simulation has made considerable strides in the last fifteen years, both as a method for understanding social processes and individual or collective behavior and as a tool for formulating and assessing social policies. Simulation consists of the creation of a computational model of social phenomena. Like other types of models, it is designed to capture key features of a system while abstracting away its nonessential features. An important development in social simulation has been that of *agent-based* social simulation (ABSS). This approach consists of constructing models of societies of artificial agents. Agents are autonomous entities with well-defined rules of behavior. Running such a model amounts to instantiating a population of agents, allowing the agents to run, and observing the interactions between them. ABSS thus differs markedly from traditional (equation-based) approaches to simulation, where relationships among conceptual entities (e.g., social groups and hierarchies, or markets and taxation systems) are expressed through a set of mathematical equations. Agent-based modeling has a number of advantages over equation-based modeling, including, notably, the ability to represent a heterogeneous population and to realistically model social networks (Axtell 2000).

Interestingly, the evolution of simulation as a means for computational study of societies has been mirrored by developments in computational modeling at the individual level. Whereas earlier models of cognition tended to emphasize one of the aspects of cognition (for instance, memory or learning), some recent approaches have been more integrative, with a focus on putting the pieces together. The products of this integrative approach are known as *cognitive architectures*, and are essentially models that capture different aspects of cognition and their interaction. Such models tend to be generic and task-independent. Cognitive architectures have greatly grown in expressive power in recent years, and now capture a variety of cognitive phenomena, including various types of memory/representation, modes of learning, and sensory-motor capabilities (see, e.g., Anderson and Lebiere 1998; Sun 2002).

So far, however, the two fields of social simulation and cognitive architectures have developed in near-isolation from each other (with some exceptions; e.g., Best and Lebiere 2003; West et al 2003; Carley and Newell 1994). Thus, most of the work in social simulation assumes very rudimentary cognition on the part of the agents (e.g., Cecconi and Parisi 1998). At the same time, while the mechanisms of individual cognition have been the subject of intensive investigation in cognitive architectures (e.g., Anderson and Lebiere 1998; Klahr et al 1987; Rumelhart and McClelland 1986; Sun 2002), the relationships between sociocultural forces and individual cognition remain largely unexplored (again with some exceptions; e.g., Hutchins 1995; Resnick et al 1991; Lave 1988).

We believe, however, that the two fields of social simulation and cognitive architectures can be profitably integrated. As has been argued before (Sun and Naveh 2004; Moss 1999; Castelfranchi 2001), social processes ultimately rest on the choices and decisions of individuals, and thus understanding the mechanisms of individual cognition can lead to better theories describing the behavior of aggregates of individuals. So far, most agent models in social simulation have been extremely simple (in the form of very simple automata with a few ad-hoc assumptions) or entirely absent (in the case of equation-based modeling). However, we believe that a more realistic cognitive agent model, incorporating realistic tendencies, inclinations and capabilities of individual cognitive agents can serve as a more realistic basis for understanding the interaction of individuals (Edmonds and Moss 2001). Although some cognitive details may ultimately prove to be irrelevant, this cannot be determined *a priori*, and thus simulations are useful in determining which aspects of cognition can be safely abstracted away.

In this paper, we first describe the model proposed by Gilbert (1997) for capturing the growth of academic science. Gilbert's model is based on the idea of artificial societies, but lacks agents capable of meaningful autonomous action. We then describe a cognitive architecture, CLARION, that captures the distinction between explicit and implicit learning. This architecture has been used to model a variety of cognitive data (see Sun and Peterson 1998; Sun et al 2001; Sun and Zhang 2004; Sun et al 2005). We demonstrate how Gilbert's simulation can be redone in an enhanced way with CLARION-based agents, and argue that the latter approach provides a more cognitively realistic basis for social simulation.

2 Gilbert's Model of Academic Science

Science develops in certain ways. In particular, it has been observed that the number of authors contributing a certain number of articles to a scientific journal follows a highly skewed distribution, corresponding to an inverse power law. This distribution, which is known as a Zipf distribution, is common to a number of other phenomena in information science, such the frequency of spoken words (Ridley and Gonzales 1994) or of links on the World Wide Web (Kleinberg 1999; Kumar et al 1999). In the case of scientific publication, the tendency of authorship to follow a Zipf distribution was observed by Lotka (1926) and is known as Lotka's law.

Simon (1957) developed a simple stochastic process for approximating Lotka's law. One of the assumptions underlying this process is that the probability that a paper will be published by an author who has published i articles is equal to a/i^k, where a is a constant of proportionality.

Using Simon's work as a starting point, Gilbert (1997) attempts to model the growth of science, including Lotka's law, through the use of an artificial society. He obtains his simulation data based on the assumption that the system selects a focal paper randomly first, which can be represented as a point in a multi-dimensional (two-dimensional) space of ideas, and then it randomly selects a number of other papers, each of which occupies a different point in the nearby region and pulls the original point in its direction slightly. The resulting paper can be located on that multi-dimensional space based on the above factors. Papers are randomly assigned authors, based on a stochastic process that takes the ratio of papers to published authors into account. To capture the constraint that academic papers must be original, a newly published paper must lie at least m coordinate units away from any other existing paper, where m is a constant.

Another assumption is that the number of papers produced in a given time period is determined by the number of papers in existence during the previous time period, by specifying a small probability of each existing paper acting as the seed for a new paper (and then by selecting an author for that paper). Thus, it is papers that spawn more papers, with authors playing only an ancillary role in the process.

Using this model, Gilbert obtained an idea space divided into clusters, which he identified as corresponding to different scientific specialties. Each cluster originated in a few seminal papers and accumulated additional papers at an increasing rate over time. This model yielded a set of publication trends that accord with human data, including the power curve described above (as can be seen in Tables 1-2). A highly uneven distribution of number of publications per author was observed, with the majority of authors publishing but one paper. A similarly skewed outcome was obtained for the number of citations received per author, with most authors receiving a modest number of citations, and a minority of authors receiving a large number of citations.

3 The Case for Cognitive Realism

Gilbert's model has the strengths common to most equation-based models: it begins with a small set of simple assumptions, and succeeds in capturing its target phenomenon (the growth of academic science) with considerable accuracy. However, to a significant extent, it is not cognitively realistic. The model assumes that authors are non-cognitive and interchangeable; it therefore neglects a host of cognitive phenomena that characterize scientific inquiry (e.g., learning, creativity, evolution of field expertise, etc.). Using a more cognitively realistic model, we can envision addressing some of these omissions, as well as exploring other emergent properties of the model and their correspondence to real-world phenomena.

The challenge, then, is to develop a model that explains macro-level phenomena in terms of micro-level cognitive processes. Such an approach would enable us to avoid the artificial assumptions often found in equation-based modeling (such as Gilbert's assumption that papers automatically spawn more papers, or that researchers are randomly assigned authorship of specific papers). In this manner, we can hope to put more distance between assumptions and outcomes, and thereby arrive at deeper explanations. This is, essentially, an instance of "docking" (Axtell et al 1996).

The next section describes a cognitive model, CLARION, that can serve as the basis for a more cognitively realistic version of Gilbert's simulation.

4 Model

4.1 Explicit vs. Implicit Learning

In an attempt to understand the processes underlying human learning, various categories of knowledge have been proposed. Among them, one enduring distinction is that between explicit and implicit—or conscious and unconscious—learning (e.g., Reber 1989; Stanley et al 1989). While both implicit and explicit learning have been actively investigated, the complex interaction between these two modes of learning has largely been downplayed or discounted (with a few exceptions; e.g., Mathews et al 1989; Sun et al 2001).

However, despite the lack of study of such interaction, recent evidence suggests that it is difficult to find a situation in which only one type of learning is employed (Reber 1989; Seger 1994). Our review of empirical data suggests that while one can manipulate conditions so that one type of learning is favored over the other, in nearly every case, both types are involved, with varying degrees of contributions from each (see, e.g., Sun et al 2001; Stanley et al 1989).

Likewise, in the development of cognitive architectures (e.g., Rosenbloom et al 1993; Anderson and Lebiere 1998), the distinction between implicit and explicit knowledge has been proposed for a long time. However, work on cognitive architectures has focused on the "top-down" direction of learning (that is, first formulating explicit rules and on that basis acquiring skills implicitly) while ignoring the "bottom-up" direction (that is, first learning implicitly and only then explicitly).

In the next subsection, we describe a cognitive architecture, CLARION, which seeks to capture the interaction between explicit and implicit learning (Sun and Peterson 1998; Sun et al 2001).

4.2 A Summary of the CLARION Architecture

CLARION is a general cognitive architecture with a dual representational structure (Sun 1997; Sun 2002). It consists of two levels: the top level encodes explicit knowledge, and the bottom level encodes implicit knowledge. See Figure 1 for a sketch of the model.

At the bottom level, the inaccessible nature of implicit knowledge is captured by a subsymbolic distributed representation provided by a backpropagation neural network. This is because representational units in a distributed environment are capable of performing tasks but are subsymbolic and generally not individually meaningful (see Rumelhart et al 1986; Sun 1994). Thus, they are relatively inaccessible. Learning at the bottom level proceeds in trial-and-error fashion, with the neural networks being guided by reinforcement learning (i.e., Q-learning; see Watkins 1989).

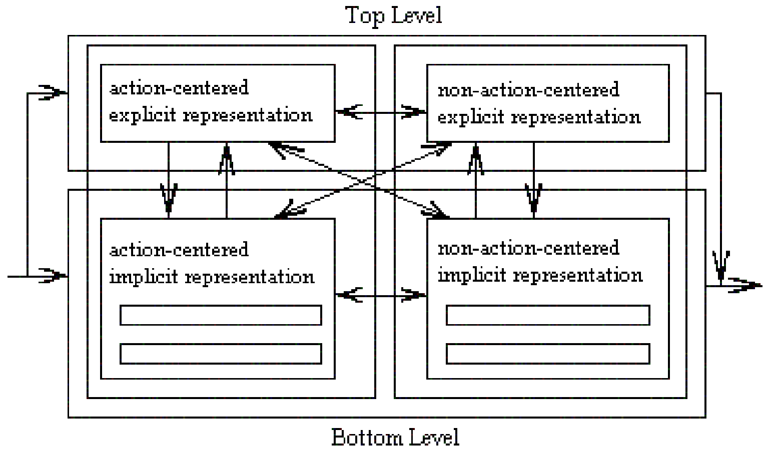

Figure 1: The CLARION architecture

At the top level, explicit knowledge is captured by a symbolic or localist representation, in which each unit is more easily interpretable and has a clearer meaning. This characteristic captures the property of explicit knowledge being more accessible and manipulable (Sun 1994). Learning at the top level involves, first, constructing a rule that corresponds to a "good" decision made by the bottom level. This rule is subsequently refined, either by generalizing or specializing it. Here, the learning process is guided by an "information gain" measure that compares the success ratio of various modifications of the current rule.

The overall algorithm of CLARION's action decision making is the following:

1. Observe the current state x.
2. Compute in the bottom level the value of each of the possible actions (a_i's) associated with the state x: $Q(x, a_1)$, $Q(x, a_2)$, ..., $Q(x, a_n)$.
3. Find out all the possible actions ($b_1, b_2, ..., b_m$) 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 (which are sent down from the top level), 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 (as will be explained later).
7. Update the top level using the *Rule-Extraction-Refinement* algorithm (explained below).
8. Go back to Step 1.

In learning settings where correct input/output mappings are known, straight backpropagation, a supervised learning algorithm, can be used to guide the neural networks. Supervised learning procedures require that a uniquely correct output be determined for each input. However, when no such input/output mappings are provided externally, reinforcement learning can be used instead, especially Q-learning (Watkins 1989) implemented in backpropagation neural networks. Such learning methods are cognitively justified. For instance, Shanks (1993) demonstrated that "associative" models (that is, backpropagation-based neural networks) do a better job of capturing human instrumental conditioning than rule-based models. Cleeremans (1997) argued that implicit learning could not be captured by symbolic models, but rather by subsymbolic ones (i.e., neural networks).

At the bottom level, a Q-value measures the "quality" of an action in a given state; that is, $Q(x, a)$ indicates how desirable action a is in state x. To acquire the Q-values, Q-learning, a reinforcement learning algorithm (Watkins 1989), is used. Q-values are then used to decide probabilistically on an action to be performed, using a Boltzmann distribution of Q-values:

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

where x is the current state, a is an action, and t controls the degree of randomness (temperature) of the process. (This method is also known as Luce's choice axiom; Watkins 1989).

Q-values are gradually tuned through successive updating, as follows:

$$\Delta Q(x, a) = \alpha(r + \gamma max_b Q(y, b) - Q(x, a)) = \alpha(r - Q(x, a))$$

where x is the current state, a is one of the actions, r is the immediate feedback, and $\gamma max_b Q(y, b)$ is set to zero for the academic science task discussed in this paper, because we rely on immediate feedback here (see details below). Using Q-learning allows reactive sequential behaviors, which rely on moment-to-moment input, to emerge.

Q-learning is implemented in backpropagation networks. Applying Q-learning, the training of the backpropagation network is based on minimizing the following error at each step:

$$err_i = \begin{cases} r - Q(x, a) & \text{if } a_i = a \\ 0 & \text{otherwise} \end{cases}$$

where i is the index for an output node representing the action a_i, and a is the action performed. Based on the above error measure, the backpropagation algorithm is applied to adjust the network's internal weights.

At the top level, explicit knowledge is captured by simple prepositional rules. An algorithm was devised for bottom-up learning—extracting rules using information culled from the bottom level, which is known as the *Rule-Extraction-Refinement*, or RER, algorithm. The basic idea is as follows: whenever an action decided by the bottom level is successful, a rule (with conditions corresponding to the current input state and an action corresponding to the one selected by the bottom level) is created and added to the top level. Then, in subsequent interactions with the world, an agent may refine a rule by considering its outcome: if successful, an agent may try to generalize a rule by relaxing its conditions to make it more universal. If a rule is unsuccessful, an agent may try to specialize a rule by imposing further constraints on the rule and making them exclusive of the current state. This is an on-line version of hypothesis testing processes studied in other contexts (e.g., Bruner et al 1956; Nosofsky et al 1994).

A rule's information gain (IG) measure is computed (in the academic science task) by considering the feedback at every step when the rule is applied. The inequality, $r > threshold_{RER}$ determines the positivity/negativity of a step and a rule governing this step (where r is the feedback received by an agent). The positivity threshold (denoted $threshold_{RER}$ above) corresponds to whether or not the agent perceives an action 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, $c1 = 1$, $c2 = 2$). This measure, essentially, compares the percentages of positive matches under different conditions A and B.

The generalization operator is based on the IG measure. Generalization involves adding an additional value to one input dimension in the condition of a rule, so that the rule will match the input more often. A rule can be generalized if the following holds

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

where C corresponds to the rule's current condition (matching the current state and action), *all* is a 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, a new rule will be generated with the condition C' possessing the highest IG measure. The generalization threshold (denoted $threshold_{GEN}$ above) determines an agent's readiness to generalize rules.

The specialization operator works in an analogous fashion, except that one value in an input dimension is deleted, rather than being added. Likewise, specialization takes place if a rule performs *worse* (rather than better) than the corresponding match-all rule. This process is described in greater detail elsewhere (Sun et al 2001). Due to running-time considerations, all simulations in this paper are run under a single (constant) specialization threshold.

To prevent useless rules from proliferating, a RER density measure is in place. A density of $1/x$ means that a rule must be invoked once per x steps or be deleted. This corresponds to the retention rate for rules, which has also been studied in humans (e.g., Hannum 1973).

To integrate the methods from the two levels, a number of methods are possible. Here, levels are selected stochastically, with a base probability of selecting each level. Other selection methods are possible as well (see Sun et al 2001).

Each level of the model may contain multiple modules, both action-centered modules and non-action-centered modules (Schacter 1990). In the current study, we focus only on the action-centered modules. The full system includes many other components (see Sun 2003 for further details).

4.3 Implementation

Both CLARION and the current simulation have been implemented as a set of Java packages. The code is available upon request. For more information, see the CLARION web page at `http://www.cogsci.rpi.edu/~rsun/clarion.html`.

Many typical psychological tasks performed by individual cognitive agents have been simulated using CLARION. The tasks include process control (PC) tasks, serial reaction times (SRT) tasks, alphabetical arithmetic (AA) tasks, and the Tower of Hanoi (TOH) task (Sun 2002). Among them, SRT and PC are typical implicit learning tasks (involving implicit reactive routines mainly), while TOH and AA are high-level cognitive skill acquisition tasks (with explicit processes). In addition, we have done extensive work on a complex minefield navigation (MN) task, which involves complex sequential decision making (Sun et al 2001, Sun and Peterson 1998). We are now in a good position to extend the modeling effort to capturing social processes. In fact, we have already worked on an organizational decision task (Sun and Naveh 2004).

5 Incorporating Cognitive Agents

In our simulation of academic science, we move to an agent-based model. Different from Gilbert's assumptions in his simulation, we treat authors as cognitive agents. Thus, authors are not merely passive placeholders, but cognitively capable individuals whose success or failure depends on their ability to learn in the scientific world. Successful authors (that is, agents who manage to identify promising research leads early on) will go on to publish numerous papers in their area, whereas unsuccessful authors will be removed from the system and replaced.

Similar to Gilbert's approach, our model characterizes the scientific world as consisting of "papers," each of which proposes a new piece of knowledge, and of authors who combine these papers to form new papers. This coincides with the constructivist view of scientific inquiry (e.g., Gergen 1985; Gergen and Gergen 1991), which sees scientific knowledge first and foremost as the product of a *constructive* process. On this view, the fruits of scientific inquiry are produced in a mutually constituting technological, linguistic, and social context. This is reflected in our model, in which papers are constructed from previous papers, and themselves serve as a basis for future constructions.

To publish a paper, an agent adopts a focal idea (as represented by an existing paper), non-randomly, in accordance with some cognitive processes. The cognitive processes may be implicit or explicit. The agent then uses other ideas (published papers) that pull the original idea in different directions (Gilbert 1997), also non-randomly, based on similar cognitive processes.

In addition, apart from utilizing existing ideas, an agent also performs local search to "optimize" the resulting idea. This reflects the fact that authors do not merely cobble together ideas from existing sources, but also try to integrate the different ideas and to refine the final product.

Because our simulation involves learning agents, there is the possibility of failure; this is important, because humans can produce papers that prove to be unpublishable. This is in contrast to Gilbert's approach, in which ideas are undifferentiated in their quality. Instead, in our simulation, each agent has a set of evaluation functions that determine the quality of ideas in the multi-dimensional idea space. These functions specify the most important considerations in terms of evaluating a scientific idea (e.g., clarity, insightfulness, empirical evidence, theoretical results, and application potential). Agents are aware of these functions. However, just as researchers in the real world cannot predict precisely when the result of their research will meet with approval and interest (Knorr-Cetina 1983), so the agents' individual valuations of these functions may differ from the community-wide valuation. This is reflected in a set of *individual* evaluation functions for each agent, consisting of a varied version of the global, or *communal*, evaluation functions.

The author population consists of CLARION-based agents. The feedback to agents is based on paper acceptance or failure. In addition, agents are provided with *partial* feedback at each step of the paper generation process, equal to a fraction of the unfinished paper's evaluation (as determined by the agent's own evaluation functions). This reflects the fact that agents do not stumble blindly through the publication process, but rather are guided to a certain extent by their experience and intuition. (Note that the partial feedback is used by the reinforcement learning algorithm, the Q-learning-Backpropagation algorithm, employed in CLARION as described earlier, which is capable of learning along the way of a publication process, based on the partial feedback, instead of waiting until the end of the publication process of a paper.)

On the one hand, an agent uses the bottom level to select a focal idea and a number of pull ideas. These two tasks are carried out by one network. The network "learns" using the *Q-learning-Backpropagation* algorithm, which corresponds to a simple form of reinforcement learning and naturally captures sequences of actions (i.e., selecting the focal idea, then the first and second pull ideas).

On the other hand, an agent uses RER rule learning to extract rules that determine: (1) how to choose focal ideas, and (2) how to choose pull ideas. These rules are used in conjunction with other rules already existing at the top level concerning local search, which represent *a priori* knowledge[1].

Agents are pre-trained for a certain number of cycles before entering the system. A cycle corresponds to a single attempt by an agent to publish a paper, whether successfully or not. Reflecting a "publish or perish" academic environment, agents are evaluated every few cycles based on their publication record (success rate). If the latter falls below a minimum expected standard, the agent is removed from the academic world. It the agent passes all the evaluations, it retires upon reaching the maximum allowable age. Whenever an agent retires (or is removed), a new agent takes its place.

5.1 Simulation Details and Parameters

Below are descriptions of the simulation parameters and details. As such, they may be skipped by readers who are not interested in the technical simulation details. The justification for the settings of these parameters either can be found in the preceding subsection, or should be self-evident.

An agent sees an idea (i.e. paper) as a multi-dimensional vector. Without loss of generality, we assume that vectors have 12 dimensions, with each dimension being a real number between 0 and 16.

The input to an agent's network for generating the focal idea (paper) is all zeroes. The input to a network for generating a pull idea is the current focal idea (as specified by a value in each dimension of the idea space), plus any pull ideas that have been selected so far[2]. An additional input is the sequential index of the current idea to be selected (0 for focal idea; 1 or 2 for pull ideas). The output of an agent's network is the current focal/pull idea, according to its numeric index in the idea space.

Pulling is accomplished by moving the original point in the space towards the second point (the "pull" point), by a certain fraction of the distance between them. To determine the amount of pull, Gilbert's formula was adopted:

$$d_i := d_i + (d_i' - d_i)(1 - m)/2$$

where i ranges over all dimensions, d_i is a dimension of the original point, d_i' is a dimension of the "pull" point, and $m \epsilon [0, 1]$ is a constant that is incremented by 0.1 after each "pull." In our simulation, we restrict agents to two "pull" ideas per focal idea.

At the end of idea selection and "pulling," an agent performs limited local search. The search is done strictly within a radius of 2 from the modified idea, using a *hill-climbing* algorithm, which has been shown to capture human reasoning in some domains (e.g., Grobman and Gilmore 2003). Hill climbing, in this case, refers to finding the best idea (according to the agent's evaluation functions) that can be obtained by increasing/decreasing one dimension of the current idea by 1, and replacing the current idea with this modified idea (if it is superior to the current idea). This process is repeated until the current idea cannot be improved without violating the radius restriction mentioned above.

The feedback to agents is based on paper acceptance (1 if a paper is accepted; 0 otherwise). Partial feedback consists of 1/3 of an unfinished paper's evaluation. [3]

To represent communal evaluation functions, five randomly-generated polynomial functions are used, with a maximum degree of three. Each agent is randomly assigned weights for these functions which they use to compute a weighted average of the five functions in evaluating ideas. Note that while agents know the "true" functions, they do not know the "true" weights for each function (i.e., the weights used by the journal in determining acceptance or rejection of an article) and therefore their learning process consists, in part, of overcoming their initial bias in this respect.

Paper acceptance is determined externally by (1) being above a threshold of 0.5 when a new paper was evaluated using a predetermined, fixed, weighted-average combination of 5 evaluation functions;

(2) having a minimum distance of 1 between a new paper and any existing papers. The former requirement sets minimum standards for paper quality, thus forcing the agent to learn and adapt. The latter requirement reflects a criterion for paper originality. [4]

Agents are trained for 10 cycles before entering the academic world. A cycle can be mapped to actual time (e.g. one year) in the life of an agent. Authors are evaluated every 5 cycles, and must demonstrate a success rate of at least 40% to be retained. There is a maximum of 10 agents living in the system at any given time, and the maximum lifespan of agents is 60 cycles. This is quite analogous to the real-life academic world.

5.2 Results

The results of our simulation are shown in Tables 1-2, along with results (reported by Simon) for *Chemical Abstracts* and *Econometrica*, and estimates obtained from previous simulations by Simon (1957) and Gilbert (1997). The figures in the tables indicate number of authors contributing to each journal, by number of papers each has published. These results are also shown graphically in Figures 2-3. The CLARION results are the averages of 300 runs, thus ensuring the representativeness of the results.

The CLARION simulation data for the two journals can be fit to the power curve $f(i) = a/i^k$, resulting in an excellent match[5]. The results of the curve fit are shown in Table 3, along with correlation and error measures. (Aggregate paper counts for authors with 11 or more papers were not included in the fit.)

Table 1: Number of authors contributing to *Chemical Abstracts*.

Number of Publications	Actual	Simon's estimate	Gilbert's simulation	CLARION simulation
1	3991	4050	4066	3803
2	1059	1160	1175	1228
3	493	522	526	637
4	287	288	302	436
5	184	179	176	245
6	131	120	122	200
7	113	86	93	154
8	85	64	63	163
9	64	49	50	55
10	65	38	45	18
11 or more	419	335	273	145

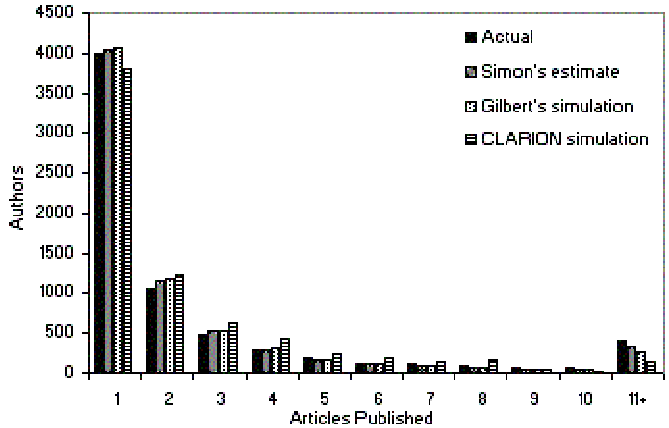

Figure 2: A graphical representation of results from Table 1.

Table 2: Number of authors contributing to *Econometrica*.

Number of Publications	Actual	Simon's estimate	Gilbert's simulation	CLARION simulation
1	436	453	458	418
2	107	119	120	135
3	61	51	51	70
4	40	27	27	48
5	14	16	17	27
6	23	11	9	22
7	6	7	7	17
8	11	5	6	18
9	1	4	4	6
10	0	3	2	2
11 or more	22	25	18	16

Table 3: Results of fitting CLARION data to a power curve.

Journal	a	k	Pearson R	R-square	RMSE
Chemical Abstracts	3806	1.63	0.999	0.998	37.62
Econometrica	418	1.64	0.999	0.999	4.15

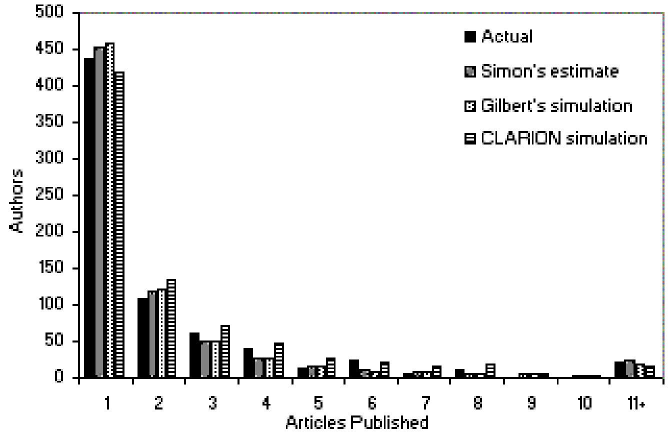

Figure 3: A graphical representation of results from Table 2.

In our simulation, the number of papers per author reflected the cognitive ability and the cognitive suitability of an author, as opposed to being based on auxiliary assumptions such as those made by Gilbert (1997). This explains, in part, the greater divergence of our results from the human data: whereas Gilbert's simulation consists of equations selected to match the human data, our approach relies on much more detailed and lower-level mechanisms—namely, a cognitive agent model that is generic rather than task-specific. The result of the CLARION-based simulation is therefore emergent, and not a result of specific and direct attempts to match the human data. We put more distance between mechanisms and outcomes, which makes it harder to obtain a match with the human data. Thus, the fact that we were able to match the human data shows the power of our cognitive agent-based approach compared to traditional methods of simulation.

6 Varying the Cognitive Parameters

In the preceding simulation, agents were run under a single set of cognitive parameters. This resulted in a Zipf distribution of articles that closely matched the human data.

However, there is more to cognitive modeling than merely replicating and validating previous results. Because CLARION captures a wide variety of cognitive phenomena, we can vary parameters that correspond to specific cognitive factors, and observe the effect on performance as function of *cognition*. This confers an important advantage over other, more task-specific models, where differences in performance tend to be artifacts of the particular model used and may be of little independent interest. With CLARION, the parameters being altered are the fundamental building blocks of cognition, and thus observed differences in performance are far more likely to stem from testable differences in individual cognition.

Accordingly, in the second simulation we vary a number of cognitive parameters and observe the effect on performance. In previous explorations with CLARION (Sun and Naveh 2004), we were able to vary all parameters simultaneously in a full factorial design. This yielded a complex pattern of interactions between the different variables tested. Here, due to the greater complexity of the task, we settle for something less—namely, each parameter is varied relative to a single baseline condition, rather than being compared to all combinations of all levels of all variables. Such an approach sacrifices some rigor, but allows a far greater range of parameters to be tested (without running into the exponentially increasing time costs of a factorial design).

Two sets of parameters of CLARION are varied. These parameters were described in detail in Section 4.2. The first set consists of fundamental properties of the model, including: (1) Learning rate of the neural networks. (2) Reliance on the top vs. the bottom level, expressed as a fixed probability of choosing each level. (3) Temperature, or degree of randomness.

The second set consists of parameters concerning RER learning and rule extraction, including: (1) RER density, which determines how often a rule needs to be encountered to avoid being deleted due to disuse. (2) RER generalization threshold, which must exceeded for a rule to be generalized.

To generate the different conditions, a baseline condition is used which is identical to the parameter set used in the first simulation. Each condition is identical to the baseline condition in all respects, except for one factor which is assigned a different value in each condition. Results for each condition consist of the average of five runs. (This was because this simulation was very time consuming, given the complexity of the cognitive architecture, as well as the number of variations that needed to be run. Thus we could only get the averages of five runs, in order to finish the simulations within a reasonable time frame. Fortunately, the variability of results was small, so the results were representative.)

6.1 Results

The effect of learning rate on performance is shown in Figure 4. An agent's learning rate essentially determines its responsiveness to success or failure. As can be seen, the best performance is obtained under a moderately high learning rate (0.1-0.3). If the learning rate is too high, an agent's recent experiences tend to disproportionately impact the learning process. This leads the agent to submit articles that are too similar to recently successful articles, which leads to more non-unique articles that are in turn rejected; equally, an agent is too swift to abandon a previously promising line of research as the result of a single rejection. If, on the other hand, the learning rate is too low, an agent will be slow to capitalize on recent successes and failures.

The fact that a balance between implicit learning and explicit rule learning is desirable can be seen in Figure 5. While increasing the reliance on implicit learning can lead to modest gains, overemphasizing it at the expense of explicit rule learning slows down the learning process dramatically. This is especially true during the initial steps of learning, when neural networks are still imperfectly trained, and rules, as crisp guidelines that are based on past success, are useful for speeding up learning.

Crucial in this connection is the existence of a high-quality rule base. This can be ensured, among other things, by the proper selection of a density measure, as shown in Figure 6. If density is too low, then rules persist even when they are no longer needed (for instance, when an agent has already exhausted a particular line of research encoded as a rule, and has moved on to other fields, represented by different combinations of ideas). On the other hand, when density is too high, even successful rules are often deleted before they can be fully utilized. In both cases, performance suffers.

An agent's exploration of the idea space is modulated, in considerable part, by the temperature (i.e., randomness) of its search process. As can be seen in Figure 7, agents are at their most prolific under a moderately high temperature setting—that is, when they show a willingness to experiment (i.e., to pursue new leads, which would not occur under a low temperature setting) while still being guided by their experience in the majority of cases. This observation accords with what we know of the role of serendipity in scientific discovery. In many areas of science—for instance, medicine—most of the major discoveries have been serendipitous, the result of seemingly irrelevant investigations. While such discoveries were formerly attributed to good fortune, it has since been argued (e.g., in Oliver 1991) that serendipity is, in part, a cognitive faculty that can be nurtured and developed. Our model captures this characteristic by adopting a modest degree of randomness in the decision-making process.

As with other parameters considered so far, an agent's generalization threshold must be carefully selected (see Figure 8). If it is set too low, even less successful rules will be generalized, leading to a lower-quality rule base. Too high, and it will prevent the generalization even of successful rules. In the latter situation, an agent will rigidly apply successful ideas only in the precise context in which they initially appeared (for instance, as the second "pull" idea in generating a paper) without recognizing their more general applicability.

As the preceding discussion shows, the cognitive parameters of individual agents are crucial in determining the rate of scientific progress in a society. By varying these parameters, we can arrive at communities that produce lesser or greater numbers of papers. Apart from this aggregate measure of scientific productivity, however, it is also interesting to see if the patterns of individual contribution observed in earlier simulations will be preserved. In particular, we want to see if the power curve from Section 5.2 is obtained under different cognitive settings. Our results show that it is. As can be seen in Figures 9-10, different settings of the density and generalization threshold

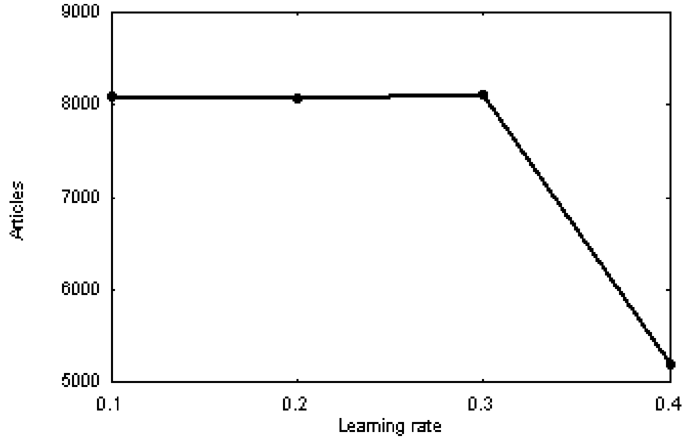

Figure 4: The effect of learning rate on collective number of articles published.

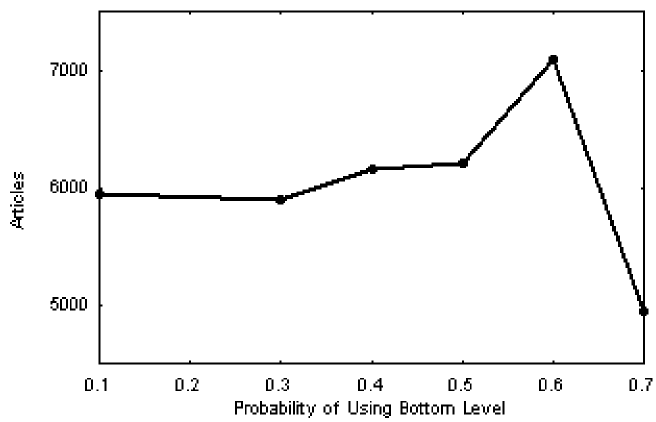

Figure 5: The effect of explicit vs. implicit learning on collective number of articles published.

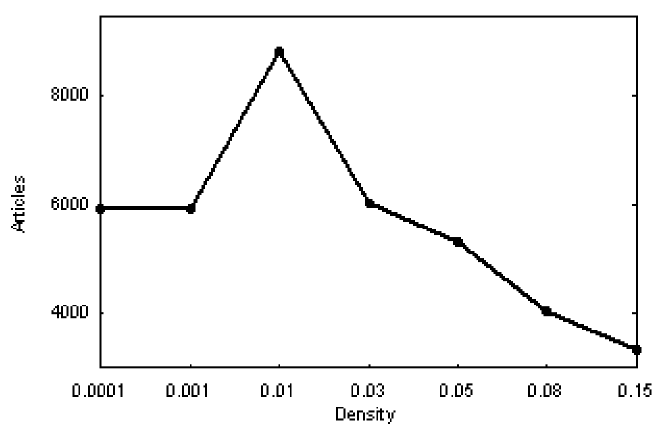

Figure 6: The effect of density on collective number of articles published.

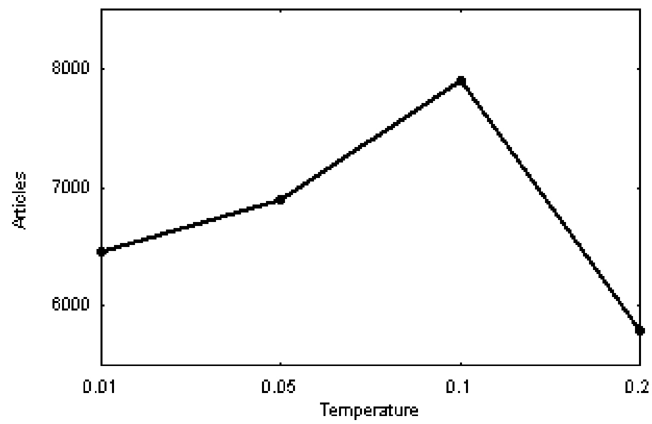

Figure 7: The effect of temperature on collective number of articles published.

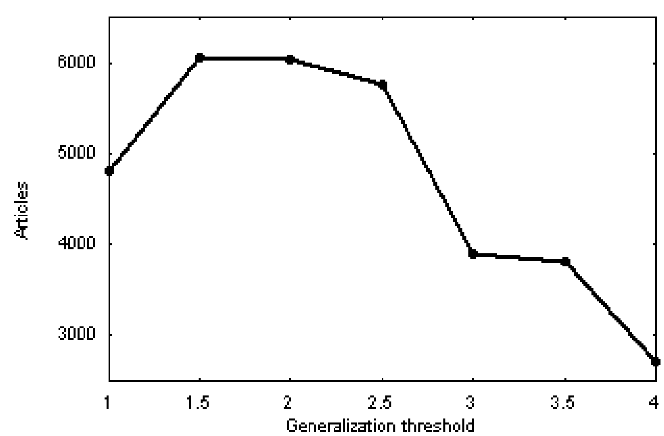

Figure 8: The effect of generalization threshold on collective number of articles published.

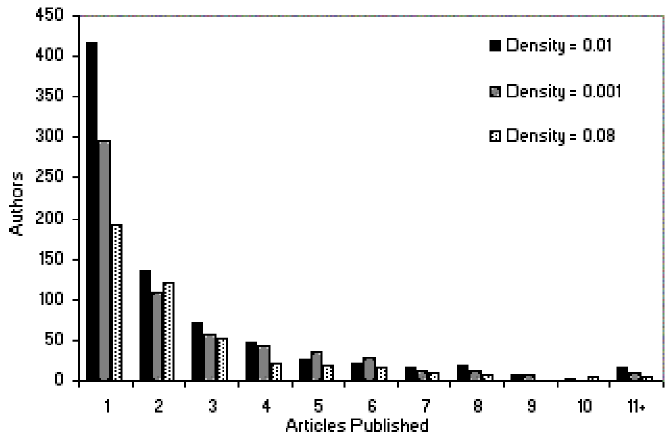

Figure 9: Authors contributing to final paper count, by number of articles that each has published. CLARION simulation results for different settings of the density parameter.

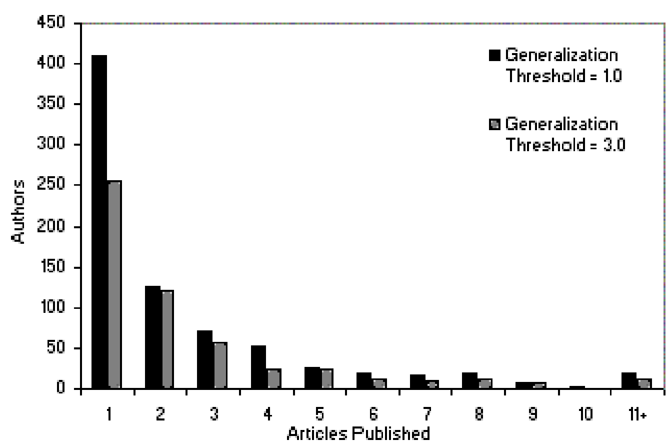

Figure 10: Authors contributing to final paper count, by number of articles that each has published. CLARION results for different settings of the generalization threshold parameter.

parameters lead to larger or smaller numbers of papers in aggregate, but they do not fundamentally change the authorship curve, which follows an inverse power distribution in all cases. Similar results were obtained for other (though not all) ranges of cognitive parameters.

This result, which may be termed cognitive-social invariance, is an important one, since it shows that some regularities that characterize societies are to some extent invariant with respect to agent cognition (within a reasonable parameter range). While some societies may prove more successful than others in terms of absolute scientific productivity, the same large-scale patterns may be observed regardless of cognitive differences. This increases our confidence in the general applicability of the model, while reducing the likelihood that the patterns observed are a byproduct of a particular set of cognitive parameters. (In contrast, in Sun and Naveh 2004, we show that some other patterns are indeed directly related to the settings of cognitive parameters).

7 Discussions

7.1 Construction in Context

Science has been characterized as a constructive, rather than descriptive endeavor (e.g., Glaserfeld 1987). This is in contrast to more traditional view of science as an investigative process that leads to the discovery of "facts" about the natural world. Instead, the constructivist interpretation regards the product of science primarily as the result of a process of fabrication. Accordingly, the process of scientific investigation is seen as leading chiefly to the construction of scientific "objects," rather than providing a factual description of nature.

Science, according to the constructivist view, is done within a highly preconstructed reality. Scientists begin with a particular theoretical conception of a field, a characteristic vocabulary, and a set of standard methodologies (Shapin 1982). The environment in which science takes place—the laboratory—is itself preconstructed and consists of particular technologies, measurement instruments, operating procedures, and so on. Moreover, the object of science is primarily a utilitarian one, and is guided by "what works." This has led some to characterize scientists as producers of technology, that is, of technical knowledge on how to manipulate things technically (rather than of knowledge that depicts nature).

This conception of researchers as producers of knowledge is reflected in our model. Much like their real-world counterparts, agents in our simulation are "practical reasoners" (Knorr-Cetina 1981) and operate by constructing new ideas from previous ones. In contrast to Gilbert's model, here authors are cognitive learners, and hence are guided by their experience in determining which combinations of resources (ideas) work and which do not. Like real researchers, agents in our model are opportunistic in seizing on and exploiting existing ideas, seeking not so much to maximize utility as to satisfice it (as embodied by an accept/reject publication criterion for papers).

Ethnographic studies of science have emphasized the contextual nature of scientific work (Knorr-Cetina 1983). When scientists are asked to explain particular choices that were made in the course of their research, they almost invariably respond by referring to the availability of a particular set of equipment, the presence of a collaborator with the technical "know-how" in a given technique, the availability of funding, and so on. In nearly every case, reference is made to aspects of the scientist's immediate experience. These aspects are so variable, personal, and wide-ranging that there is no hope of reducing them to a small set of criteria or principles that can be used to predict a scientist's behavior. Instead, some sociologists have contented themselves with characterizing researchers as contextually-embedded deciders, and pointing to the vast importance of a scientist's individual circumstances, and their interaction, in determining the course of research.

Among the contextual aspects that constrain a scientist's decisions is the social element (Gergen 1985). Scientists operate within a linguistic and conceptual reality created through a process of social exchange. Furthermore, knowledge, as one of many coordinated human activities, "is subject to the same processes that characterize any human interaction (e.g., communication, negotiation, conflict, rhetoric)" (Schwandt, p. 127). It is for this reason that scientists have been characterized as "socially situated reasoners" (e.g., Haraway 1988; Knorr-Cetina 1981). Some of these social aspects appear to be more or less cooperative—for instance, scientific communities. Such communities are typically devoted to a research specialty, and typically involve collaboration and interaction between their members. Other, less "cooperative" accounts have stressed the competitive aspects of interaction between scientists. Among these are quasi-economic models that see the scientific field as a community of specialists competing for creative currency in the form of credibility or recognition. Both the "cooperative" and the "competitive" aspects of science are represented in our model. In CLARION-based simulations, the former occurs when agents build on ideas formulated by other researchers, thereby contributing to the idea space themselves. The latter occurs when agents preempt each other in exploiting a good combination of ideas, effectively "stealing" papers from each other.

Certainly, we should not leave out the mutual constitution of agents and contexts—the activities of one scientist not only are constrained by contexts, but also help to create further contexts and constraining further work in a field. Our simulations indeed embody such a process.

7.2 Inferential vs. Holistic Changes

Traditionally, scientific thought has been characterized as a process of linear logical inference. On this view, inferences are done as a series of steps, each with one or more conclusions that follow from a set of premises. Such an inference might look like this: as a ship appears on the horizon, one sees first its mast and then later the deck. It is as if a ship is coming over a hill. But this phenomenon is observed everywhere. Therefore the earth must be hill-shaped everywhere. Therefore the earth is round.

However, an alternative view of scientific inference has been proposed by Thagard (1992; see also Kuhn 1962). Instead of the serial process described above, he suggested that scientific theories can be evaluated in a more holistic way, on the basis of their coherence. This entails weighing postulates of different theories in parallel, accepting some and rejecting others in a way that maximizes their coherence. The notion of coherence has been used to model various historical shifts in scientific thinking, for instance the triumph of a wave-based account of light over a particle-based one in the nineteenth century (Eliasmith and Thagard 1997).

In CLARION, both of the above forms of reasoning are represented at the individual level. It has been suggested that the holistic nature of coherence-based reasoning can best be modeled by a connectionist paradigm (involving neural networks; Eliasmith and Thagard 1997). In CLARION, this is represented by learning at the bottom level. At the same time, the more traditional form of linear, explicit reasoning is captured by the top level. Thus, by using CLARION to model agents in simulations, we can account for the role of both forms of reasoning in the development of scientific theories.

7.3 Cognitive and Social Realism

One important aim of this study has been to determine whether the results of a previous model of academic science (Gilbert 1997) can be reproduced without resorting to the broad simplifications of an equation-based model. The results of our simulation suggest that the observed growth of academic science can indeed be captured even if one migrates to an agent-based model. Such a migration offers several important benefits. First, it allows us to leave behind certain artificial assumptions (for instance, that papers automatically spawn more papers). Second, it affords us the opportunity of studying the macro-level repercussions of behavior at the micro level. This was seen in Section 6, where we were able to vary parameters of explicit versus implicit learning, and measure the effect on communal performance. Third, it allows us to study patterns of interaction between individual agents. Although the latter interactions occur only indirectly in our model (either through the collision of too similar papers generated by different authors, or through the exploitation of others' ideas in generating new papers) they nonetheless result in a model that is more socially realistic than Gilbert's equation-based model.

What makes our approach unusual, however, is not that it represents actors as agents, but that it takes agent cognition seriously. So far, most agent models in simulations have been rather simple, with little attention being paid to the mechanisms of individual cognition. This study shows that a more cognitively realistic simulation, with CLARION, can replicate the results of earlier simulations. It thus provides a dual corroboration of these models, by showing them to be independent both of whether or not an agent-based model is used, and of whether or not cognitive representations are involved. Therefore, while some cognitive details clearly cannot be abstracted away, others can, and along the way, we discover important cognitive-social invariances.

Apart from validation, however, cognitive realism in simulations may lead us to better representations of the target phenomenon. For instance, in Section 6.1, we identify a possible way of representing the role of serendipity in science: namely, as a researcher's willingness to explore apparently suboptimal combinations of ideas, rather than adhering to "tried-and-true" sequences. The ability to represent such aspects of observed phenomena in terms organic to the agent model, rather than through auxiliary assumptions (for instance, by adding a "randomizing" function to the idea selection process in Gilbert's simulation) is an advantage of cognitively realistic simulations.

Another advantage of cognitive realism is that it allows us to theorize about the relative role that individual cognitive factors play in the emergence of large-scale social phenomena. Thus, in Section 6.1, we were able to vary parameters of CLARION that corresponded to aspects of cognition and tested their effect on outcomes. Our investigation showed, for instance, that the tendency to engage in inductive reasoning (that is, an agent's generalization threshold) could dramatically influence the number of papers generated by the community. It moreover suggested that this phenomenon could be described by a u-shaped curve. Such results suggest how patterns of communal thinking may change as a consequence of shifts at the individual cognitive level.

Conversely, cognitively realistic simulations may help us establish the constancy of some observed phenomena. As seen in Section 6.1, the same power curve for authorship was observed under different cognitive settings, even when overall scientific productivity differed for these conditions. Such observations may lend support to theories of cognitive-social invariance, or the independence of social phenomena from particular cognitive characteristics. Alternatively, they may suggest boundary conditions under which such phenomena begin to break down. In either case, there is considerable scope for enriching the predictions obtained from simulations.

7.4 Classifying Actors in Simulations

So far, we have discussed the benefits of embedding cognitive agents in social simulations. At the heart of this approach lies a particular conception of how *actors* should be modeled in simulations. To evaluate this approach, we must therefore compare it with other ways of representing actors in simulations. The first alternative approach, the *equation-based* approach, involves abstracting the actors away altogether. Actors in such simulations are not explicitly represented as part of the system, and their functionality is only indirectly captured by the equations. A second approach involves representing actors through the use of *agents* (i.e., autonomous computational entities). Such an approach lacks the simplicity and elegance of an equation-based approach, but its greater accessibility allows it to be evaluated and critiqued by a wider range of researchers. Moreover, using agents enables, in many cases, a more detailed representation of the target phenomenon.

Simulations vary widely in the level of detail of their actors, ranging from very simple models, such as those used in some simulations of the prisoner's dilemma (Axelrod 1987) to more detailed models, such as Anderson's ACT-R (Anderson and Lebiere 1998). Actors in simulations can be further distinguished based on their computational complexity (as expressed by computer scientists using "Big-O" and similar notations). Such measures have important implications with respect to a model's scalability, since they determine whether its running time and space requirements vary linearly (or polynomially, exponentially, etc.) with the size of its inputs. Finally, simulations differ in the degree of rationality imputed to their actors. Some simulations (for instance, in traditional economics) assume perfectly rational actors, whereas others consist of boundedly rational actors that aim merely for satisficing solutions, rather than optimal ones.

The above distinctions lead us to a set of dimensions, or axes, for classifying simulations according to their representation of actors. These dimensions include, first, whether or not a model is agent-based; second, the granularity, or detailedness of the model; third, the model's computational complexity; fourth, whether rationality is bounded or unbounded in the model; and fifth, the degree of cognitive realism in the model. These dimensions may be correlated to some extent in actuality, but they should be separately evaluated nevertheless, in order for us to gain a better picture of the relative position of a model with regard to other existing or potential models. In particular, the final dimension is seldom used in the evaluation of simulations, but it is important, for reasons mentioned in the previous subsection.

By referring to this classificatory system, we can arrive at sets of limitations common to certain classes of simulations. For instance, using the dimensions above, we can categorize our CLARION simulation as an agent-based simulation, reasonably detailed, computationally complex, boundedly rational, and cognitively realistic. This simulation therefore inherits the limitations associated with each of these five characteristics. Thus, as a high-granularity model, CLARION can make it hard to disentangle the respective contributions of different factors to the results of simulations. Likewise, its relatively high computational complexity can raise issues of scalability. Bounded rationality may hinder our ability to generalize from the results of simulations (Ahlert 2003). Finally, the choice of a cognitively realistic agent model may itself rest on particular ontological conceptions of the target phenomenon, and thus the CLARION model may not be appropriate for all simulations.

8 Conclusions

While interest in simulation as a way of describing social phenomena continues to grow apace, the issue of choosing a realistic cognitive model has not been sufficiently emphasized overall. We propose using more complex cognitive models, known as cognitive architectures, to capture human behavior. Our model of scientific publication assumed that, in constructing new ideas from previous ones, authors were guided chiefly by cognitive processes. By paying more attention to the details of individual cognition, we can arrive at more accurate representations of target social phenomena. We can also learn which cognitive mechanisms are significant in shaping social interactions, and which are not. Finally, we can study the emergence of large-scale social behavior from micro-level cognitive processes.

Acknowledgments

Xi Zhang deserves thanks for implementing the CLARION system. Thanks are also due to Jay Hutchinson for implementing a preliminary version of the simulation

Notes

[1]This knowledge is assumed to have been acquired through experience in other situations.

[2]Both pull and focal ideas are selected from an initial space of 30 ideas.

[3] We tested minor variations of the feedback. The simulation results were robust with regard to minor variations in the feedback setting.

[4] We tested variations of the above settings. The simulation results were robust with regard to variations in the above settings.

[5]Note that compared with Simon's data, there are slightly more authors who contribute a moderate (2-8) number of articles. Thus, the curve descends somewhat less steeply than the curves for the human data.

References

Ahlert, M. (2003), "An Axiomatic Approach to Bounded Rationality in Negotiations," paper read at the 2003 European Meeting of the Economic Science Association, September 18-21, Erfurt, Germany.

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

Axelrod, R. (1987), "The Evolution of Strategies in the Iterated Prisoner's Dilemma," in L. Davis (Ed.) Genetic Algorithms and Simulated Annealing, Pitman.

Axtell, R.L. (2000), "Why agents? On the varied motivations for agent computing in the social sciences," Proceedings of the Workshop on Agent Simulation: Applications, Models and Tools, Argonne National Laboratory, IL.

Axtell, R.L., J. Axelrod and M. Cohen (1996), "Aligning Simulation Models: A Case Study and Results," Computational and Mathematical Organization Theory, 1(2), 123-141.

Best, B.J. and C. Lebiere (2003), "Teamwork, Communication, and Planning in ACT-R: Agents Engaging in Urban Combat in Virtual Environments," Proceedings of the IJCAI 2003 Workshop on Cognitive Modeling of Agents and Multi-Agent Interactions, Acapulco, Mexico.

Bowers, K., G. Regehr, C. Balthazard and K. Parker (1990), "Intuition in the context of discovery," Cognitive Psychology, 17, 72-110.

Bruner, J., J. Goodnow and J. Austin (1956), *A Study of Thinking.* Wiley, NY.

Carley, K.M. and A. Newell (1994), "The Nature of the Social Agent," Journal of Mathematical Sociology, 19(4), 221-262.

Castelfranchi, C. (2001), "The Theory of Social Functions: Challenges for Computational Social Science and Multi-Agent Learning," Cognitive Systems Research, special issue on the multidisciplinary studies of multi-agent learning, 2(1), 5-38.

Cecconi, F. and D. Parisi (1998), "Individual Versus Social Survival Strategies," Journal of Artificial Societies and Social Simulation, 1(2). Available online at http://www.soc.surrey.ac.uk/JASSS/1/2/1.html.

Citro, C.F. and E.A. Hanushek (Eds.) (1991), *Improving Information for Social Policy Decisions: The Use of Microsimulation Modeling. Vol I: Review and Recommendations.* National Academy Press, Washington, DC.

Cleeremans, A. (1997), "Principles for Implicit Learning," In D. Berry (Ed.) How Implicit is Implicit Learning?, Oxford University Press, Oxford, UK.

Conte, R. and N. Gilbert (1995), "Introduction," in G. N. Gilbert and R. Conte (Eds.) Artificial Societies: The Computer Simulation of Social Life, UCL Press, London, UK.

Edmonds, B. and S. Moss (2001), "The Importance of Representing Cognitive Processes in Multi-Agent Models," in G. Dorffner, H. Bischof, and K. Hornik (Eds.) Artificial Neural Networks—ICANN'2001, Springer-Verlag: Lecture Notes in Computer Science, 2130, 759-766.

Epstein, J.M. and R.L. Axtell (1996), *Growing Artificial Societies: Social Science from the Bottom Up.* MIT Press, Cambridge.

Gergen, K.J. (1985), "The Social Constructionist Movement in Modern Psychology," American Psychologist, 40(3), 266-275.

Gergen, K.J. and M.M. Gergen (1991), "Towards Reflexive Methodologies," in F. Steier (Ed.) Research and Reflexivity. Sage, Newsbury Park, CA.

Gilbert, N. (1997), "A simulation of the structure of academic science," Sociological Research Online, 2(2). Available online at
http://www.socresonline.org.uk/socresonline/2/2/3.html.

Glaserfeld, E. (1987), *The Construction of Knowledge.* Intersystems Publications, Seaside.

Grobman, K.H. and R.O. Gilmore (2003), "The gradual emergence of domain-general problem solving strategies during infancy," manuscript under revision for Child Development.

Hannum, W.H. (1973), *A Study of Select Factors Influencing the Retention of Rules.* Ph.D. Thesis, Florida State University, Florida.

Haraway D. (1988), "The Science Question in Feminism as a Site of Discourse on the Privilege of Partial Perspective," Feminist Studies, 14(3), 575-579.

Hutchins, E. (1995), "How a Cockpit Remembers Its Speeds," Cognitive Science, 19, 265-288.

Karmiloff-Smith, A. (1986), "From Meta-Processes to Conscious Access: Evidence from Children's Metalinguistic and Repair Data," Cognition, 23, 95-147.

Klahr, D., P. Langley and R. Neches (Eds.) (1987), *Production System Models of Learning and Development.* MIT Press, Cambridge.

Kleinberg J.M., R. Kumar, R. Prabhakar, R. Sridhar, and A.S. Tomkins (1999), "The Web as a Graph: Measurements, Models, and Methods," Proceedings of the 5th International Conference on Computing and Combinatorics, July 26-28, Tokyo.

Knorr-Cetina, K.D. (1981), *The Manufacture of Knowledge.* Pergamon Press, Oxford.

Knorr-Cetina, K.D. (1983), "Towards a Constructivist Interpretation of Science," in K.D. Knorr-Cetina and M. Mulkay (Eds.) Science Observed, Sage, CA.

Kuhn, T. (1962), *The Structure of Scientific Revolutions.* Chicago University Press, Chicago.

Kumar, S.R., P. Raghavan, S. Rajagopalan, and A. Tomkins (1999), "Trawling the Web for Emerging Cyber-Communities," Proceedings of the 8th WWW Conference.

Lave, J. (1988), *Cognition in Practice.* Cambridge University Press, Cambridge, UK.

Lotka, A.J. (1926), "The Frequency Distribution of Scientific Productivity," Journal of the Washington Academy of Sciences, 16, 317-323.

Mandler, J. (1992), "How to Build a Baby," Psychology Review, 99(4), 587-604.

Mathews, R., R. Buss, W. Stanley, F. Blanchard-Fields, J. Cho and B. Druhan (1989), "Role of Implicit and Explicit Processes in Learning From Examples: a Synergistic Effect," Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 1083-1100.

Moss, S. (1999), "Relevance, Realism and Rigour: A Third Way for Social and Economic Research," CPM Report No. 99-56. Center for Policy Analysis, Manchester Metropolitan University, Manchester, UK.

Oliver, J.E. (1991), *The Incomplete Guide to the Art of Discovery.* Columbia University Press, NY.

Reber, A. (1989), "Implicit Learning and Tacit Knowledge," Journal of Experimental Psychology: General, 118(3), 219-235.

Resnick, L.B., J.M. Levine, and S.D. Teasley (1991), *Perspectives on Socially Shared Cognition.* American Psychological Association, Hyattsville, MD.

Ridley, D.R. and E. Gonzales (1994), "Zipf's Law Extended to Small Samples of Adult Speech," Perceptual and Motor Skills, 79, 153-154.

Rosenbloom, P., J. Laird and A. Newell (1993), *The SOAR Papers: Research on Integrated Intelligence.* MIT Press, Cambridge.

Rumelhart, D. and J. McClelland (Eds.) (1986), *Parallel Distributed Processing I.* MIT Press, Cambridge.

Schacter, D. (1990), "Toward a Cognitive Neuropsychology of Awareness: Implicit Knowledge and Anosagnosia," Journal of Clinical and Experimental Neuropsychology, 12(1), 155-178.

Schwandt, T.A. (1994), "Constructivist, Interpretivist Approaches to Human Inquiry," in N.K. Denzin and Y.S. Lincoln (Eds.) Handbook of Qualitative Research, Sage, Newbury Park, CA.

Seger, C. (1994), "Implicit Learning," Psychological Bulletin, 115(2), 163-196.

Shanks, D. (1993), "Human Instrumental Learning: A Critical Review of Data and Theory," British Journal of Psychology, 84, 319-354.

Shapin, S. (1982), "History of Science and its Social Reconstructions," History of Science, 20, 157-211.

Simon, H.A. (1957), *Models of Man, Social and Rational.* Wiley, NY.

Stanley, W., R. Mathews, R. Buss and S. Kotler-Cope (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), 553-577.

Sun, R. (1994), *Integrating Rules and Connectionism for Robust Commonsense Reasoning.* Wiley, NY.

Sun, R. (1997), "Learning, Action, and Consciousness: A Hybrid Approach Towards Modeling Consciousness," Neural Networks, special issue on consciousness, 10(7), 1317-1331.

Sun, R. (2002), *Duality of the Mind.* Lawrence Erlbaum Associates, Mahwah, NJ.

Sun, R. (2003), *A Tutorial on CLARION 5.0.* Available online at
`http://www.cogsci.rpi.edu/~rsun/sun.tutorial.pdf`

Sun, R., E. Merrill and T. Peterson (2001), "From Implicit Skills to Explicit Knowledge: A Bottom-Up Model of Skill Learning," Cognitive Science, 25(2), 203-244.

Sun, R. and I. Naveh (2004), "Simulating Organizational Decision-Making Using a Cognitively Realistic Agent Model," Journal of Artificial Societies and Social Simulation, 7(3). http://jasss.soc.surrey.ac.uk/7/3/5.h

Sun, R. and T. Peterson (1998), "Autonomous Learning of Sequential Tasks: Experiments and Analyses," IEEE Transactions on Neural Networks, 9(6), 1217-1234.

Sun, R., P. Slusarz and C. Terry (2005), "The Interaction of the Explicit and the Implicit in Skill Learning: A Dual-Process Approach," Psychological Review. 112(1), 159-192.

Sun, R. and X. Zhang (2004), "Top-Down Versus Bottom-Up Learning in Cognitive Skill Acquisition," Cognitive Systems Research, 5(1), 63-89.

Thagard, P. (1992), *Conceptual Revolutions.* Princeton University Press, Princeton.

Watkins, C. (1989), *Learning with Delayed Rewards.* Ph.D. Thesis, Cambridge University, Cambridge, UK.

West, R.L., C. Lebiere and D.J. Bothell (2003), "Cognitive Architectures, Game Playing, and Interactive Agents," Proceedings of the IJCAI 2003 Workshop on Cognitive Modeling of Agents and Multi-Agent Interactions (Ron Sun, Ed.), Acapulco, Mexico.