Computational and Mathematical Organization Theory (2019) 25:4–23
https://doi.org/10.1007/s10588-018-09291-0

S.I.: SBP-BRIMS 2018

Multi-scale resolution of neural, cognitive and social systems

Mark G. Orr[1] · **Christian Lebiere**[2] · **Andrea Stocco**[3] · **Peter Pirolli**[4] · **Bianica Pires**[5] · **William G. Kennedy**[6]

Published online: 24 February 2019

Abstract

We recently put forth a thesis, the *Resolution Thesis*, that suggests that cognitive science and generative social science are interdependent and should thus be mutually informative. The thesis invokes a paradigm, the reciprocal constraints paradigm, that was designed to leverage the interdependence between the social and cognitive levels of scale for the purpose of building cognitive and social simulations with better resolution. We review our thesis here, provide the current research context, address a set of issues with the thesis, and provide some parting thoughts to provoke discussion. We see this work as an initial step to motivate both social and cognitive sciences in a new direction, one that represents unity of purpose, an interdependence of theory and methods, and a call for the careful development of new approaches for understanding human social systems, broadly construed.

Keywords Cognitive modeling · Agent-based modeling · Social simulation · Multi-scale systems

1 Introduction

The degree of overlap between cognitive science and generative social science is small despite a shared interest in human behavior and a reliance on computer simulation. The former focuses, largely, on developing computational and formal accounts

The research is (partially) based upon work supported by the Defense Advanced Research Projects Agency (DARPA), via the Air Force Research Laboratory (AFRL). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, the AFRL or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. This article is an extended version of Orr et al. (2018).

✉ Mark G. Orr
mo6xj@virginia.edu

Extended author information available on the last page of the article

of human thought, action, performance and behavior with non-trivial incorporation of neurophysiological principles when warranted. The latter approaches the question of understanding social structure and dynamics using computational and formal accounts that implement simple agents (what we call *sans cognitive*) in social contexts. We submit that the dearth of interdisciplinary work between these disciplines does not serve either well. Our central thesis, the *Resolution Thesis*, is this: *the correct resolution of both cognitive and social systems depends on mutual constraints between them in the sense that the dynamics and structure of one system should inform the theoretical nature of the other*. We mean this in the context of theory development and related applications in both cognitive science and generative social science. The method implied by this thesis is what we call the reciprocal constraints paradigm—a bi-directional dependence across levels of scale w.r.t. their respective parameter specifications.[1]

Our thesis implies two claims. First, cognitive models should be able to match and predict the real world dynamics of social systems when embedded in social simulation, and, if not, the cognitive model should be questioned. Second, if an agent-based simulation is not informed by cognitive first principles, it will fail to generalize its account of the dynamics of social system to new situations.

In what follows, we will (1) flesh out the details of the reciprocal constraints paradigm, (2) review some prior work that is directly relevant to our thesis and puts it in context of recent research, (3) address issues and their potential mitigation, and, (4) close with some brief, but potentially provocative suggestions. We deliberately exercised a narrow focus using the ACT-R cognitive architecture as our primary vehicle of rhetoric, partly because it reflects our expertise, and partly because this architecture is comparatively well suited for integration of both neural and social constraints. Cognitive architectures, as opposed to any cognitive model, have the potential to capture what agents, in the scheme of generative social science, are supposed to do—make adaptive decisions that affect the environment.

2 The reciprocal constraints paradigm

Figure 1 captures the core components of the reciprocal constraints paradigm: multiple levels of scale, multiple potentials for model types at each level, and, the constraints among levels. To understand the paradigm, it will be useful to imagine a potential implementation. Consider a modeling problem in which there is a simple social system (e.g., a multi-player repeated economic game). The cognitive model is developed, with some consideration for key neural processes, call this $CM^{[1]}$, and without direct comparison to newly generated individual-level data sources (e.g., running single-subject experiments in pseudo-game like contexts). $CM^{[1]}$ is then implemented in a social network graph that controls information flow (e.g., knowing past decisions of other players) and, given some other parameterizations, a

[1] Because cognitive systems are sometimes tightly yoked to neurophysiology, we consider three levels as central to our thesis: neurophysiology, cognitive architecture, and social systems

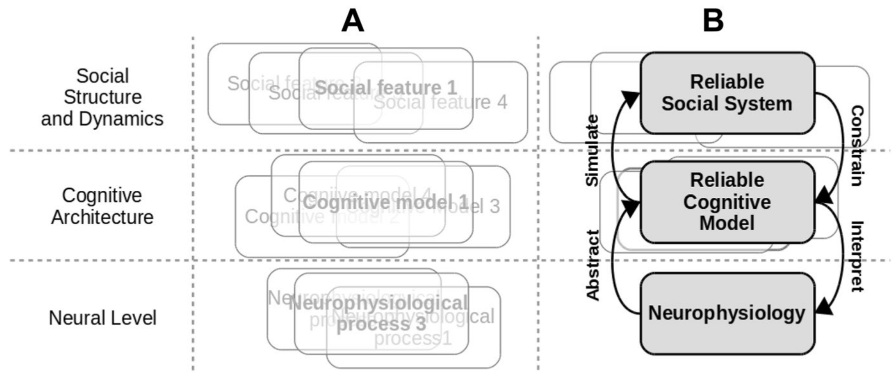


Fig. 1 The Architecture and Implementation of the Reciprocal Constraints Paradigm. Each row represents a level of scale (as labeled in the left-most column). Column A is notational for the degree of variety of potential types of neural processes and cognitive models that could be constructed to capture a phenomenon and the types of features in the social space (e.g., peer-network)—i.e., it captures the feature/model space of a particular implementation. Column B shows the implementation of the reciprocal constraints paradigm; each arrow represents a kind of constraint: *Abstract*—abstraction of neural processes to cognitive processes; *Simulate*—simulating social systems in which humans behavior is defined as a cognitive architecture; *Constrain*—the feedback signal from the accuracy of the social simulation w.r.t. to empirical measurements on human systems; and, *Interpret*—refinement of the selection of neural processes that are implicated in the cognitive model. The former two constraints we call *upward constraints*; the latter are called *downward constraints*. Implementation of the paradigm will require iteration between the feature/model space and the simulation of social and cognitive models. There may be potential for automation of this paradigm once it is well developed

simulation of the multi-player repeated game is conducted; call this SM[1]. Then, SM[1] data is aggregated in some way isomorphic to human data in a similar experimental paradigm and an accuracy/error/confidence metric is computed, call it Constraint[1] to map onto Fig. 1. (Notice, at this point, the only direct comparison to human data was at the social system level.) Constraint[1] would then be used—in an undefined way at this point—to change some aspect of the cognitive model, either directly within the cognitive level or, potentially, through the neurophysiological level. Let's imagine that it makes sense to consider neurophysiological processes as the next step, a step we call Interpret[1] to map onto Fig. 1. Now, a set of targeted neurophysiological measurements are captured by running single-subject experiments in pseudo-game like contexts which yields insight into a potential missing abstraction of neurophysiological process in the cognitive model, which we call Abstraction[1]. The cognitive model is then refactored to incorporate Abstraction[1] and the process is repeated by another simulation using the next generation of the cognitive model CM[2]. Note, this example provides only one of an infinite set of paths; the paths may be consequential to the final model and could include integration of human data at one or more points.

A fundamental part of the paradigm is the acknowledgment that scaling up from the cognitive level to the social level is different, in principle, compared to the scaling up from the neural to cognitive level. The former transition instantiates multiple isomorphic and interdependent cognitive models as a simulated system (i.e., each model is functionally equivalent). The latter, in contrast, abstracts information processing functionalities that are assumed to be interdependent but potentially different in nature (i.e., not functionally equivalent). This is an important difference in light of what a constraint actually means.

3 Illustrative prior work

We do not know of any examples of work that directly implement the *Reciprocal Constraints Paradigm*—a principled use of constraining information between levels of scale, e.g., considering both the *Simulate* and *Constrain* arrows between the cognitive and social levels of scale in Fig. 1. However, precursors exist. For example, from a largely cognitive perspective, there are several relevant threads of work that address aspects which are important for the *Resolution Thesis*, e.g., on multi-agent systems (Sun 2006), computational organizational theory (Prietula et al. 1998), computational social psychology (Vallacher et al. 2017), work practices (Sierhuis et al. 2007) and Anderson's *Relevance Thesis* (Anderson 2002). These efforts, in the main, attempted to provide both more accurate predictions of social system level behavior and explanations of such that were grounded in cognitive first principles. Neurophysiological considerations have, as we will illustrate below, followed a similar pattern in terms of informing higher levels of scale (cognitive) from the lower neurophysiological level of scale. Next, we describe efforts that either inform ACT-R from neurophysiology or use implementations of ACT-R as the agent definitions in a social simulation all of which are considering constraints from lower to higher levels of scale. Further, we offer a glimpse of how generative social science has conceptualized the integration of cognitive first principles into the behavior of agents to date.

3.1 The ACT-R cognitive architecture

Computational modeling aims to quantitatively capture human cognitive abilities in a principled manner. Cognitive architectures are computational instantiations of unified theories of cognition that specify the structures, representations and mechanisms of the human mind. Cognitive models of any given task can be developed using a cognitive architecture as a principled implementation platform constraining performance to the powers and limitations of human cognition. Cognitive models are not normative but represent Simons (1991) theory of bounded rationality (Simon 1991), and can also represent individual differences in knowledge and capacity such as working memory. Cognitive models can be used to generate quantitative predictions in any field of human endeavor.

ACT-R is a highly modular cognitive architecture, composed of a number of modules (e.g., working memory, procedural and declarative memory, perception and action) that operate in parallel asynchronously through capacity-limited buffer interfaces. Each module in turn consists of a number of independent mechanisms, typically including symbolic information processing structures combined with equations that represent specific phenomena and regularities (e.g., power law of practice and forgetting, reinforcement learning). Most notably, the architecture includes a number of learning mechanisms to adapt its processing to the structure of the environment. ACT-R has been applied to model human behavior across a wide range of applications (see ACT-R web site for over a thousand publications), ranging from basic experimental psychology paradigms to language, complex decision making, and rich dynamic task environments. The combination of powerful computational mechanisms and human capacity limitations (e.g., working memory, attention, etc.) provides a principled account of both human information processing capabilities as well as cognitive biases and limitations.

3.2 Neurophysiogical constraints in ACT-R

The development of ACT-R has been guided and informed, in recent years, by the increased understanding of the computational mechanisms of the brain. For example, independent modules have been associated to specific brain regions and circuits, and this correspondence has been validated multiple times through fMRI experiments. The detailed computations of crucial ACT-R components can also be derived from the neural mechanisms they abstract. For instance, the latency to retrieve declarative information from long-term memory can be derived from the dynamics of the integrate-and-fire neural model (Anderson 2007), and the mechanisms for skill acquisition can be derived from reinforcement learning (Anderson 2007) as well as from the simulation of the large-scale effects of dopamine release in the fronto-striatal circuits (Stocco et al. 2010). In fact, the modularity of ACT-R permits to easily abstract and integrate lower-level neural principles within the architecture. While this approach does not grant the full flexibility of large-scale neural simulations, it has been repeatedly shown to be very effective in capturing features of human behavior that would otherwise have remained unexplained, while at the same time maintaining the computational parsimony of a cognitive symbolic architecture. For example, implementing the dynamics of memory retrieval permits to capture a variety of decision-making effects and paradoxes, beyond those explained by current mathematical models (Gonzalez et al. 2003). The modularity of ACT-R also permits to regulate the degree of fidelity of a module to its biological counterpart, without affecting the entire architecture. As an example, Stocco (2018) has shown that the competition between the direct and indirect pathways of the basal ganglia can be captured by splitting production rules into opposing pairs. This procedure captures the cognitive effects of Parkinsons disease, and provides a way to model individual differences in decision-making (Stocco 2018) and cognitive control (Stocco et al. 2017) that are due to individual differences in dopamine receptors in the two pathways. See Fig. 2.) This is an example of additional mechanisms that can be added to ACT-R to incorporate further biological details (i.e., the *abstract* constraint in Fig. 1).

3.3 Social simulation with ACT-R agents

To study the dynamics of simple systems, work using ACT-R has focused on iterated two-player games, including both adversarial games (e.g., paper-rock-scissors, pitcher-batter in baseball) and social dilemmas allowing both cooperation and competition dynamics such as Prisoners Dilemma and Chicken Game West

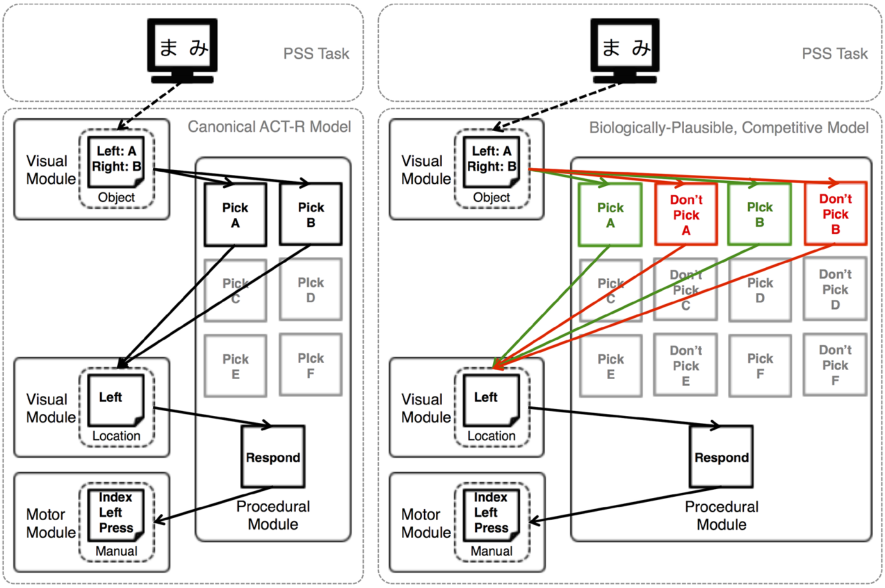


Fig. 2 An example (taken from Stocco (2018) with permission) of how neurobiological constraints can be incorporated in a cognitive architecture. The two panels illustrate two alternative ways to implement a forced choice task with six possible options (A through F) in ACT-R. (Left Panel) A canonical ACT-R model, in which each option A...F is associated with a single, corresponding production rule (Pick A Pick F). In this model, the expected value of the different options is encoded as the expected utility of each production rule. The utility of each rule is learned through reinforcement learning in ACT-Rs procedural module, which is associated with the basal ganglia. However, the lack of biological plausibility in ACT-Rs procedural module prevents the model from capturing the results of the original study. (Right Panel) A biologically-plausible version of the same model, in each of the original production rules is split into two opposite actions (Pick A Pick F and Dont Pick A Dont Pick F), whose utilities are learned separately. This new version abstracts the competition between the direct and indirect pathways of the basal ganglia circuit. When equipped with this biologically-plausible version of production rules, the model can successfully reproduce the results in the neuropsychological literature, as well as capture individual differences in genetics (Stocco 2018) and even correctly predict new findings (Stocco et al. 2017)

and (Lebiere 2001; Lebiere et al. 2003, 2000). Even in such simple systems, we have observed the emergence of complex effects such as bifurcations and stochastic resonance (West et al. 2005). To scale up to more complex yet regular systems, we have modeled the emergence of group consensus and choice differentiation in networks of a few dozen nodes on tasks such as consensus voting and map coloring, respectively, and observed phenomena such as sensitivity to network rewiring parameters (Romero and Lebiere 2014). To study complex cognition in complex systems, we have designed and implemented an information foraging task called the Geogame that involves cooperative and competitive problem solving and have observed effects including sensitivity to network topology and tradeoffs between perceptual and memory strategies (Reitter and Lebiere 2012). Clearly, this work represents well the *simulate* constraint in Fig. 1.

A common pattern in models of social interaction using ACT-R has been to ground agent decisions in previous experiences, whether explicitly in the form of memories or implicitly by reinforcement of existing strategies, as mentioned in the previous section. We will focus here on two examples using the former approach, because it has been both more common and more flexible. Models of adversarial interaction usually involve a core capability of detecting patterns in the opponent behavior and exploiting them until they disappear. For instance, playing paper rock scissors involve exploiting the human limitation in generating purely random behavior (the standard game theory solution) by detecting statistical patterns in move sequences. An expectation of an opponents next move can be generated by matching his most recent moves against previous sequences using statistical memory mechanisms. Once a pattern is being exploited, the opponent is likely to move away from it and in turn exhibit new ones, requiring a cognitive system that constantly unlearns previous patterns and learns emerging ones, rather than traditional machine learning systems that are training on a fixed set of inputs and then frozen. In that sense, social simulation is the ultimate challenge for real-time learning agents: unlike physical environments which change relatively slowly and can be mastered in a relatively static way, social interactions (especially competitive and adversarial interactions), as they involve other cognitive entities, are endlessly evolving and require constant learning and adaptivity.

Another example of this approach of decisions from experience in the social domain is our model of event generation in the GitHub web repository. The instance-based learning modeling methodology (Gonzalez et al. 2003) assumes that experience is represented in the form of memories associating a context representation to the decisions taken in that situation. In our GitHub specific model, the decisions involve generating events related to three cryptocurrency repositories (repos for short). The context of each decision is the daily price changes of the three corresponding cryptocurrencies: Bitcoin (BTC), Ethereum (ETH) and Monero (XMR). The decision itself consists in selecting a specific user, repo, type of event (e.g., pull request), and event time.

One notable aspect of this model is that it does not represent an individual decision maker, as is typical of cognitive models, but instead stands for a community of users. A clustering algorithm was used to generate compact communities. Figure 3 shows some of the social structure inherent to GitHub[2]. The models memory was loaded with the events involving all the users in a given community (typically about a thousand). Events are generated through retrievals from memory performed in multiple steps (see Fig. 4). At each step, the most active event in memory is retrieved. The activation of an event reflects its recency through a process of power law decay, and the degree to which it matches the specified context. Initially, the context is simply the daily price change of the three cryptocurrencies. The partial matching mechanism decrements the activation of a chunk to the extent that it mismatches the context. That will tend to favor events having occurred in similar

[2] We don't have an approximation of the degree to which our community clustering algorithm mapped onto the structure shown in Fig. 3.

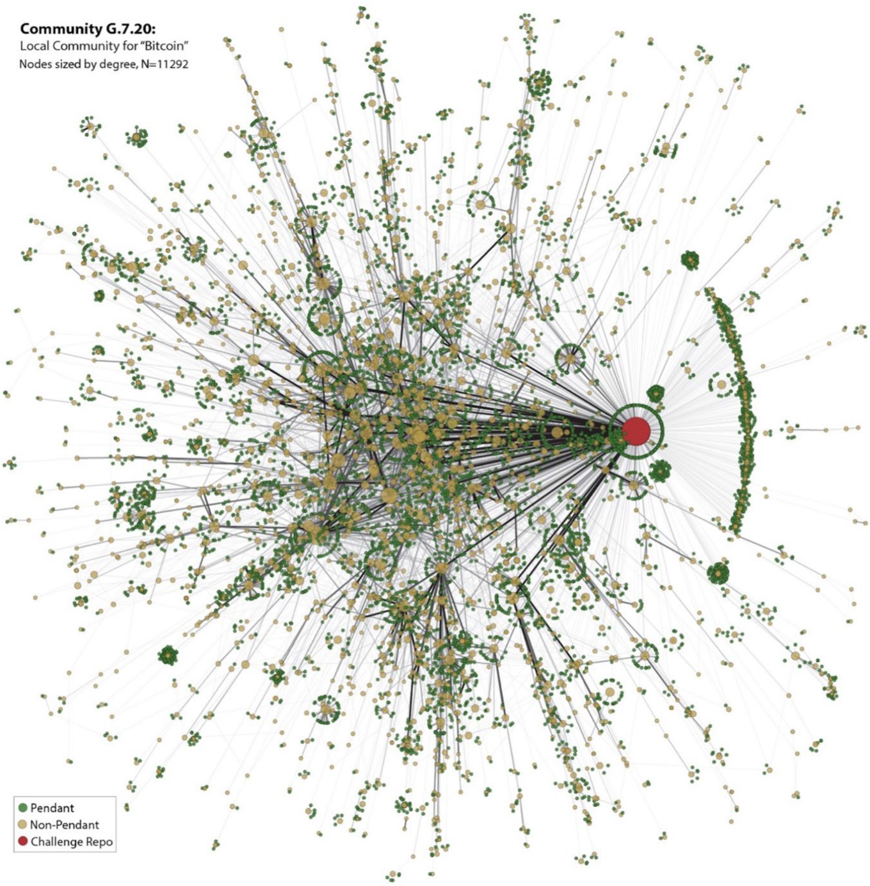
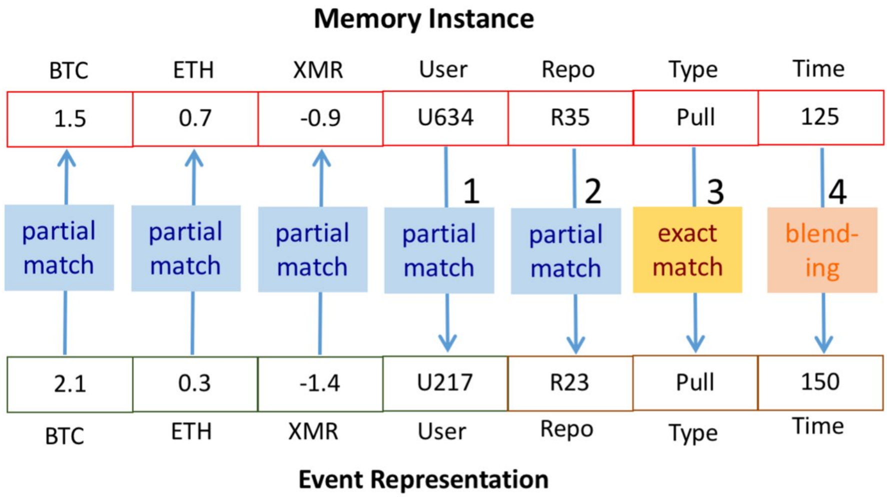


Fig. 3 Bitcoin GitHub Community Graph. Community G7.20. This community is extracted from the larger GitHub interaction network (n = 54.1M nodes, m = 134.16M edges). Ties represent any interaction from a user to a repo—although this is a bipartite graph, the node attribute (user/repo) is not depicted, except for the identification of the BitCoin repo. Clusters were identified in two stages via repeated application of the Louvain Community detection routine (Blondel et al. 2008), where we first identify large superclusters and then clustered any with more than 50,000 nodes again to identify sub-clusters. Supercluster 7 is of size n = 269,264, m = 385,315, subcluster 20 has n = 11,292 m = 14,197). Layout is based on a Fruchterman–Reingold algorithm, with degree = 1 nodes (green) placed in a circle near their respective hub (tan). The Bitcoin Repo is indicated in red and offset slightly to highlight its position in the network. (Color figure online)

circumstances (i.e., market conditions) in the past. The first retrieval results in a user in the community being selected, reflecting the probability of that user generating events in the (recent) past. That user is then added to the context before the second retrieval to select a repo, likely one that that user has most commonly interacted with, although it could also be a repo that a similar user has interacted with. The repo is then added to the context to select an event type, largely reflecting the distribution of events between that user and repo (or similar ones). Finally, the time



$$V = \underset{V_j}{\operatorname{argmin}} \sum_{i=1}^{k} P_i \times Sim(V_j, v_{ij})^2 \qquad B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right)$$

$$M_i = A_i + \sum_{j=1}^{l} MP \times Sim(d_j, v_{ij}) \qquad A_i = B_i + \sum_{j=1}^{m} W_j \times S_{ji}$$

Fig. 4 ACT-R memory schematic for simulating GitHub agents. See text for details

of the event is generated by applying the blending mechanism to a fourth retrieval to average the distribution of time intervals involving similar events. The process results in a model that generates a stream of events that reflects past distributions of events while allowing for some generalization to similar events through the memory retrieval mechanisms.

3.4 Comparison to the generative social science approach

Generative social simulation has historically been concerned with the simulation of interacting agents according to simple behavioral rules. We can often equate the outcome behavior of agents to a simple binary action (e.g., you either riot or dont riot) and the behavioral rules that produce this outcome to simple mathematical and logical formulations (e.g., if/else statements, threshold values). We are in debt to the many classic models that made computational social science the field it is today (Schelling 1969; Axelrod et al. 1995; Epstein 2002). However, there has been some acknowledgment that to gain further insight into social systems, we need to decompose behavior into its underlying cognitive, emotional, and social (interactions) processes. With this, we are beginning to see a slight shift to developing models with more complex agents (Epstein 2014).

In this vein, an approach that has gained some traction is the use of conceptual frameworks that integrate the varied components of agent decision-making processes (Caillou et al. 2017; Sakellariou et al. 2008; Malleson et al. 2012; Pires and Crooks 2017). Such frameworks include BDI (Beliefs, Desires, and Intentions) and PECS (Physical conditions, Emotional state, Cognitive capabilities, and Social status) (Kennedy 2012). In the BDI framework, beliefs are said to be the individuals knowledge about the environment, desires contain information about the priorities and payoffs associated with the current objective, and intentions represent the chosen course of action Rao et al. (1995). BDI agents use a decision tree process which relies on payoff and utility maximizing functions to select goals and to determine the optimal action sequence for which to achieve those goals. The focus on optimality, however, may pose limits on its ability to model the boundedly rational agent and has been criticized for being too restrictive (Rao et al. 1995). PECS views agents as a psychosomatic unit with cognitive capabilities residing in a social environment (Schmidt 2000). The PECS framework is flexible due to its ability to model a full spectrum of behaviors, from simple stimulus-response behaviors to more intricate reflective behaviors, which requires a construction of self that necessitates the agent be fully aware of its internal model. By example, Pires and Crooks Pires and Crooks (2017) used the PECS framework to guide implementation of the underlying processes behind the decision to riot, applying theory from social psychology to create the agents internal model and to simulate social influence processes that heightened certain emotions and drove the agent's towards certain actions. These frameworks, while helpful for guiding implementation, are not to be considered substitutes for cognitive architectures such as ACT-R. They can, however, provide a meta-framework [sometimes called a macro-architecture) to organize knowledge and skill content in respect to a cognitive architecture (e.g., (West et al. 2017)].

Cognitive architectures and meta-frameworks are fundamentally complementary (Lebiere and Best 2009). Cognitive architectures precisely specify the basic cognitive acts that can be used to compose complex models in a bottom up approach, but provide few constraints to guide those complex structures. Meta-frameworks provide a top down methodology to decompose complex tasks into simpler ones and structure the knowledge required, but do not include a principled grounding for that process. The combination of the two approaches can be achieved in a number of different ways. One approach is to develop integrated environments allowing modelers to flexibly leverage the two methodologies in a way that is best suited to the specific requirements of each application (Lebiere et al. 2008). An alternative is to provide high-level patterns and abstractions that can be formally compiled into cognitive models in a target cognitive architecture (Ritter et al. 2012).

4 Issues and their mitigation

4.1 Downward constraints

4.1.1 Social to cognitive

This issue was laid out plain by Allen Newell about three decades ago (Newell 1990) in reference to the *social band* (corresponding to time intervals > 10^4 seconds that represent organizational behavior and other social systems). Newell, thinking in terms of the strength of a system's levels, hypothesized that social bands should be characterized as having weak strength, and therefore may not be computing much at all, in a systematic way. If Newell's surmises are correct, then constraining cognitive architectures from the social band makes little sense. Anderson's Relevance Thesis (Anderson 2002), put forth about a decade later, does not address the operation of social systems in terms of constraining cognitive models; his thesis is more focused on the degree to which understanding lower bands, especially the cognitive (10^{-1} to 10^1 time scale), are implicated in qualities of higher bands, e.g., educational outcomes. So, from the cognitive perspective, there might not be much signal from the social band that could serve as a useful constraint on cognitive architectures.

However, there are potential approaches towards mitigation of this problem, Newell's thesis notwithstanding. Online social communities often exhibit emergent empirical regularities. For instance, the World Wide Web exhibits many regularities including the small world organization of link structure and the distribution of the lengths of browsing paths that users exhibit. The latter has been called the Law of Surfing. Many of these regularities have been modeled at the social level using variants of statistical mechanics. The Law of Surfing (Huberman et al. 1998) observes that the frequency distribution of path lengths (number of Web pages visited) is well fit by an Inverse Gaussian Distribution, that has a long positive tail. The key insight at the social level is that a Web surfer can be viewed as moving around in a kind of space analogous to the Brownian motion of a small particle on a liquid surface. In the case of the Web surfer, the movement is in the dimension of expected utility that will be received (or not) when visiting a Web page, where the expected utility from continuing on to the next page is stochastically related to the expected utility of the current page, and the Web surfer continues until a threshold expected utility is reached. This is modeled as a stochastic Wiener process. But, the Law of Surfing can also be predicted from Monte Carlo simulations with ACT-R agents (Fu and Pirolli 2007). In contrast to the stochastic social models, these finer-grained ACT-R agents can make predictions for specific Web tasks at specific Web sites, which can be used to predict and engineer improvements (Chi et al. 2003). However, the emergence of the Law of Surfing from the ACT-R agent simulations is seen as constraint on the cognitive models.

In short, the social band, at least in some domains, does have structure that could constraint cognitive modeling efforts. A question that remains is to what degree will it be possible to develop general methods across the varieties of social domains for the purpose of constraining cognitive models.

4.1.2 Cognitive to neurophysiology

The downward Interpret arrow in Fig. 1 could seem paradoxical, given that the underlying neural level is often taken as the ground truth of the entire system. Neurophysiological findings, however, are often only imperfectly understood. For instance, the existence of basal ganglia projections outside of the frontal lobe was considered impossible for a long time until recently (Middleton and Strick 1996). Even when our grasp of neurophysiology is solid, cognitive architectures can be helpful in providing a functional interpretation to existing data by focusing on the computational integration of different circuits, that is, answering the question of what does this circuit do?. The most famous example in this sense is the interpretation of the activity of dopamine neurons in terms of reward prediction error signals in reinforcement learning (RL) (Schultz et al. 1997)—an interpretation that borrowed from a decades-old AI theory [temporal difference learning: Sutton (1988)] to solve decades of seemingly inconsistent empirical findings on the role of dopamine (Bunney et al. 1991; Schultz 2002). Incidentally, this example perfectly illustrates how the Interpretation is further aided by the use of a comprehensive architecture on an agents behavior, such as that provided by RL agents. In our case, the adoption of a single architecture (such as ACT-R) to create multiple models provides the unifying framework to interpret neurophysiological data. The fact that the activity of the same neuronal process must be interpreted in the same way across multiple models of different tasks provides additional constraints to maintain the interpretation consistent.

4.2 Upward constraints

4.2.1 Parsimony and generative social science

By uncovering some new relationship or testing some stylized hypothesis of social phenomena many classic agent-based models [e.g., (Schelling 1969; Epstein 2002)] have demonstrated the value of modeling simple (*sans cognitive*) agents. For instance, Reynolds (1987) illustrates how three simple rules of behaviors can result in the emergence of the collective behavior of a flock of birds—what looks like the highly coordinated actions of a “leader” is actually the result of three simple rules.[3] These models and many others in the computational social sciences adhere to parsimony, or keeping the model simple such that the model has just enough of right features and no more, as a main guiding principle (Miller and Page 2009). Arguments for this approach stress the intuitive and interpretive appeal of such models (Miller and Page 2009; Gigerenzer et al. 1999). The purpose of the model may also dictate that the model be parsimonious [e.g., (Reynolds 1987)]. In short, parsimony in respect to simple agents has served well as a strategy in generative social science. It

[3] Agent-based models, however, can range in abstraction, from the stylized models just described to empirically-driven models; although the latter in no way implies incorporation of cognitive constraints.

is natural, then, to ask if cognitive modeling breaks with this notion of parsimony in modeling social systems.

We think the issue of parsimony in generative social science does not imply anything particular about the use of cognitive architectures in social simulations. Parsimony implies that model simplicity is considered in conjunction with how well a model matches empirical findings. Thus, the issue of whether to include cognitive agents, as defined in the reciprocal constraints paradigm, is largely an empirical issue. We offer that cognitive constraints may provide the right model and thus improve the degree to which a social simulation matches empirical findings. Moreover, because cognitive models inherit mechanistic constraints from cognitive architectures, they might actually end up being more parsimonious than agent-based models without such constraints.

4.3 Implementing at-scale, multi-scale, theory-heavy, data-heavy social simulations

The *ResolutionThesis* claims that at-scale social simulations[4] will be improved, in terms of a close comparison to real-world empirical data, when informed by neural and cognitive sciences. To date, there is very little simulation work that is both at-scale and multi-scale in a sociological, cognitive and neural sense. This may, in part, stem from the fact that these types of simulations present unique challenges (with compute resources, theory integration, data fusion, etc.).

Towards the goal of developing at-scale, multi-scale, theory-heavy, data-heavy social simulations, we have recently developed a simulation platform called the *Matrix Agent-Based Modeling Platform*[5], shown in Fig. 5. In terms of the *Reciprocal Constraints Paridigm* a few points are relevant: (1) the Matrix was developed with the precise objective of ingesting modeling efforts in the cogntive science and artificial intelligence communities that are designed to capture some form of human agency. Thus, it is agnostic to the programming language used to define agents. One can write agent models that work with the Matrix in any programming language, e.g, Python, R, C and more specialized applications coming from the cognitive science and artificial intelligence communities. The model code interacts with the Matrix platform using programming language agnostic interprocess commmunication (IPC) methods (like JSON over TCP/IP); (2) When using the Matrix platform, model authors can use generic data structures to afford implementation of virtually any social structure (e.g., social interaction graphs) and support the needs of existing cogntive architecture libraries (like ACT-R) and deep neural networks (like Lens, PyTorch, or TensorFlow); (3) Scalability is built into the platform itself such that given a heavy workload (e.g., having large numbers of agents), the computation can

[4] Simuilations on the order of say 10^4 to 10^7 agents; examples might include large communities within social media platforms; a bipartite graph among patients and care providers within a national health system; daily interactions in a multi-national corporation.

[5] The ACT-R GitHub example described above used this platform

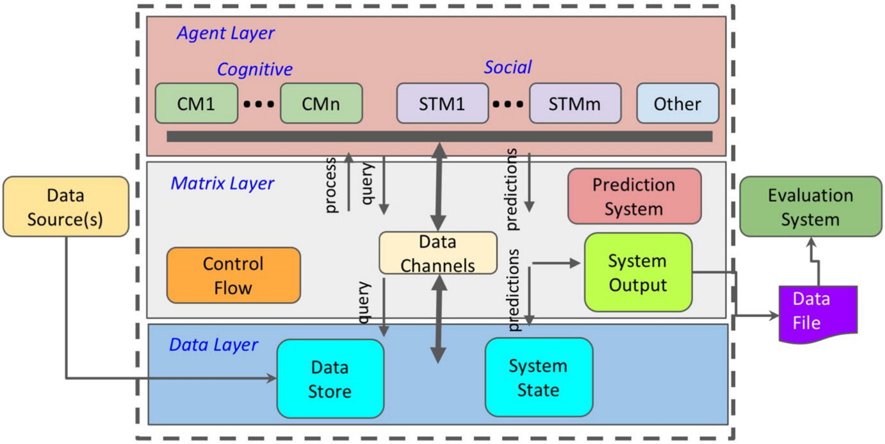


Fig. 5 The matrix agent-based modeling platform: a distributed software environment for agent-based modeling and simulation (implemented and developed in Python). It is responsible for (i) updating and providing a system view to agents, (ii) control flow logic, (iii) orchestrating agents, (iv) message passing, where (collections of) agents are independent software processes. Agents are programming language independent. Agent developers focus on serial implementation of their algorithms

be seamlessly distributed across large numbers of machines, without any change in the model code.

Through the deveolpment of the *Matrix* and some preliminary at-scale implementations with same, we now possess some useful observations, the principal of which is that parameter optimization is complicated by the fact that it may be intertwined with neural, cognitive and social theory. There are related implications, some of which are as follows: First, and something that may preclude the development of the class of at-scale social simulations we are considering here, is the fact that much of the theory we need to parameterize may not be in usable form (computational or mathematical) or may not exist at all. For example, across a large swath of psychology, the theory is represented as an interpretation of the statistical analysis of experimental paradigms without benefit of a formal mathematical or computational model of proceeses or representational forms. In a similar way, larger-scale population observations in sociology and demography, for example, that are based on complex survey design only allow inference at the behavioral level and not the psychological or cognitive level. Second, simulation parameters may be deeply tied to theory and, as such, were likely tuned by other methods (e.g., designed experimental results with human subjects). So, what is the precise nature of parameter optimization? Does it, for example, include rerunning the experimental paradigms with human subjects (in the case of psychological/cognitive theory)? If yes, then the parameters of the experimental paradigm become, in essence, part of the parameter optimization problem. And, how would we develop experimental methods with the right level ecological and task validitiy for a given social simulation goal? When is it valid to use purely observational data? Finally, in the limit, a set of simulation parameters that capture a set of theoretical assumptions might be thrown out completely. What would be used as a substitute? Another theory or something non-theoretic?

One approach we have used with the *Matrix* to deal with the challenges of parameter optimization leverages a methodology that (Reitter and Lebiere 2010) formulated called *accountable modeling*. That approach is not only a technical solution to scaling up the cognitive architecture but also a scientific commitment to an approach that explicitly states which aspects of the model are constrained by the architecture and which are free parameters to be estimated from data. This commitment helps determine which aspects of the social-scale simulation reflect the cognitive mechanisms and can be assumed to generalize, and which have been parameterized to reflect aspects of the situation not constrained by first principles, and thus will need to be estimated from data in new situations. Such an approach actually results in simpler, more transparent models that are explicit about their parameters rather than trying to camouflage them under a mechanistic veneer.

Given the varied issues around parameter optimization, one approach we are implementing currently is the use of highly modular agent architectures. This approach builds a task-based behavioral model of agents that captures the statistics coming from observational data (e.g., Twitter, Reddit or Github behaviors) in a way that is, in the main, devoid of social, cognitive or neural theory. This is considered a base model from which aspects of neural, cognitive and social theory can be integrated in a modular way.

Our observations wrought from our experience in building an at-scale, multi-scale, data-heavy, theory-heavy social simulation really amount to a set of questions and issues, with less in the way of answers—the *Matrix* was developed to begin their exploration. It is probably obvious, but social simulations of the type we are considering here will require sustained efforts that are deeply interdisciplinary in nature.

4.4 Language

The study of how languages change over time (e.g., from Old English of Beowulf through the Middle English of Chaucer's Canterbury Tales to modern-day English) leads to a view in which language is a fundamentally fluid social construction that arises and changes due to complex interactions between the communal need for effective communication and the cognitive constraints on learning and processing in individual language. In the cultural evolutionary framework (Croft 2008), the emergence of increasingly regular and complex utterances arises through the processes of (1) the continual generation of novel elements and constructions by individual language users, and (2) the selection, reuse, and propagation of variants by other users as a function of cognitive biases and social/communication influences (Hruschka et al. 2009). This framework is, in concept, within the domain of the *RCP*.

There are two key dimensions in studying language change: the social contact structure among language learners (agents) and the language acquisition mechanism invoked for them. In the literature, contact structure is most commonly of two classes: random mixing, whereby contact is assumed to be equally likely among members of a population, and iterated, whereby each member of the population teaches only one other member of the population in a one-dimensional chain. Recent work has begun to examine the implications of more complicated social network structure (Ke et al. 2008; Lou-Magnuson and Onnis 2018). However, a clear limitation of most existing work is the lack of sophistication in the assumptions about diffusion dynamics in the context of realistic, human-like social network structure. Further, the mechanisms for individual language acquisition range from the very simple and abstract, e.g., replicator dynamics (Nowak et al. 2001; Pais et al. 2012) and simple contagion (Fagyal et al. 2010) to Bayesian inference (Griffiths and Kalish 2007) to detailed psychological process models (e.g., neural networks capturing phonology/morphology at a micro level (Hare and Elman 1995; Polinsky and Van Everbroeck 2003). The advantage of simple language acquisition mechanisms when dealing with populations is that the dynamics of language evolution can often be studied analytically, but only at the cost of psychological plausibility.

Human language, we submit, promises to be a very fruitful area for development of the *RCP* and the exploration of *Resolution Thesis*. Our understanding of language processing at the cognitive and psychological level is precise (in terms of temporal scale), developed in computational formalisms and understood from a developmental perspective. We can paint a similar, parallel story with respect to our understanding of social structure and dyamics. The combination of these two in a principled way should yeild benefit.

5 Discussion

Crossing levels of scale or analysis inevitably takes one near to deep scientific issues that echo notions pointed out 50+ years ago in Simon's "Architecture of Complexity" paper (Simon 1962) [see also (Anderson 1972) for similar early example]. Our thesis does not imply that the simplifying assumptions used to study large scale social or economic phenomena (using economic or sociological methods) are wrong, but that they are, in the end, deeply intertwined with human cognitive and neural processes. Nor does our thesis claim that, because cognitive and neural models were not developed with large scale social systems in mind, they are wrong. We do, however, want to point out that one should acknowledge that when studying a part of a social system (e.g., cognition) it is nontrivial to get the other levels of scale right (e.g., social network dynamics). It is equally important, when designing at-scale simulations, to remember that the academic and professional disciplines that are called upon to contribute theory will mostly likely not provide off-the-shelf functionality in terms of parameters and functions.

In some sense, what we know about social systems is a patchwork of many disciplines and methods. The *Reciprocal Constraints Paradigm* was designed to provide a tighter knit, so to speak. In this sense, its a repair, but one with a clear objective. The *RCP* attempts to provide what is naturally given to the study of biological structures in nature, e.g,. the macrotermes bellacosi nest (temite mound), where a key objective is to understand how are such structures built from a multi-scale perspective: genetic, neurobiological, cognitive, behavioral and social. Each perspective lends a hand in explaining the structure of the mound. More generally, the study of the system and its constitutent features and parts is driven by a set of explananda (e.g., the termite mound: how did it get there and what does it do?). This is the perspective and approach advocated by the *Reciprocal Constraints Paridigm* for studying human social systems—the constraints are designed to get us, potentially, to the same goal as if we used a more holistic biologically-inspired approach from the start.

Beyond the *RCP*, we might reasonably consider starting from scratch by leveraging the perspective of biology[6] for the study and simulation of human social systems. The driving explanandum might start at the top (e.g., markets, sexual reproduction patterns, social stratification, warfare, government structures, etc.) and work would then continute both down and up levels of scale as needed to explain key macro and micro patterns of the social system, but with the key explananda in mind. This notion isn't meant as a mere provocation, but only as a way to guage what we already know and to provide potential alternatives to knowing it differently. These and other questions fall beyond our thesis and the *RCP* but may provide fodder for future generations and help to keep in perspective the current state of the art in terms of our understanding human social systems. The decision to replace an old quilt, patches and all, with something entirely new is never easy, and not only for sentimental reasons. Some of the patches were undoubtedly well constructed and, potentially, still of use.

Acknowledgements The *Matrix Agent-Based Modeling Platform* was developed by Parantapa Bhattacharya, Saliya Ekanayake, and Mandy Wilson.

References

Anderson JR (2002) Spanning seven orders of magnitude: a challenge for cognitive modeling. Cognit Sci 26(1):85–112

Anderson JR (2007) How can the human mind occur in the physical universe?. Oxford University Press, Oxford

Anderson PW (1972) More is different: broken symmetry and the nature of the hierarchical structure of science. Science 177(4047):393–396

Axelrod R (1995) A model of the emergence of new political actors. The computer simulation of social life. Artificial societies, London, pp 19–39

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech 10:P10008

Bunney BS, Chiodo LA, Grace AA (1991) Midbrain dopamine system electrophysiological functioning: a review and new hypothesis. Synapse 9(2):79–94

Caillou P, Gaudou B, Grignard A, Truong CQ, Taillandier P (2017) A simple-to-use bdi architecture for agent-based modeling and simulation. In: Advances in Social Simulation 2015, Springer, New York, pp 15–28

Chi EH, Rosien A, Suppattanasiri G, Williams A, Royer C, Chow C, Cousins S (2003) The bloodhound project: automating discovery of web usability issues using the infoscent simulator. In: ACM conference on human factors in computing systems, pp 505–512

Croft W (2008) Evolutionary linguistics. Annu Rev Anthropol 37:219–234

Epstein JM (2002) Modeling civil violence: an agent-based computational approach. Proc Natl Acad Sci USA 99(3):7243–7250

[6] This idea has obvious roots in Wilson's Sociobiological approach. Wilson (2000)

Epstein JM (2014) Agent_Zero: toward neurocognitive foundations for generative social science. Princeton University Press, Princeton
Fagyal Z, Swarup S, Escobar AM, Gasser L, Lakkaraju K (2010) Centers and peripheries: network roles in language change. Lingua 120(8):2061–2079
Fu WT, Pirolli P (2007) Snif-act: a model of user navigation on the world wide web. Hum Comput Interact 22(4):355–412
Gigerenzer G, Todd PM (1999) Simple heuristics that make us smart. Oxford University Press, Oxford
Gonzalez C, Lerch FJ, Lebiere C (2003) Instance-based learning in dynamic decision making. Cognit Sci 27(4):591–635
Griffiths TL, Kalish ML (2007) Language evolution by iterated learning with bayesian agents. Cognit Sci 31(3):441–480
Hare M, Elman JL (1995) Learning and morphological change. Cognition 56(1):61–98
Hruschka DJ, Christiansen MH, Blythe RA, Croft W, Heggarty P, Mufwene SS, Pierrehumbert JB, Poplack S (2009) Building social cognitive models of language change. Trends Cognit Sci 13(11):464–469
Huberman BA, Pirolli P, Pitkow JE, Lukose RM (1998) Strong regularities in world wide web surfing. Science 280(5360):95–97
Ke J, Gong T, Wang WS (2008) Language change and social networks. Commun Comput Phys 3(4):935–949
Kennedy WG (2012) Modelling human behaviour in agent-based models. In: Agent-based models of geographical systems, Springer, New York, pp 167–179
Lebiere C, Best BJ (2009) From microcognition to macrocognition: architectural support for adversarial behavior. J Cognit Eng Decis Making 3(2):176–193
Lebiere C, Wallach D, West R (2000) A memory-based account of the prisoners dilemma and other 2 x 2 games. In: Proceedings of International Conference on Cognitive Modeling, Universal Press, Kansas, pp 185–193
Lebiere C, Gray R, Salvucci D, West R (2003) Choice and learning under uncertainty: A case study in baseball batting. In: Proceedings of the 25th Annual Meeting of the Cognitive Science Society, Mahwah, Erlbaum, pp 704–709
Lebiere C, Archer R, Best B, Schunk D (2008) Modeling pilot performance with an integrated task network and cognitive architecture approach. Hum Perform Model Aviat
Lou-Magnuson M, Onnis L (2018) Social network limits language complexity. Cognit Sci
Malleson N, See L, Evans A, Heppenstall A (2012) Implementing comprehensive offender behaviour in a realistic agent-based model of burglary. Simulation 88(1):50–71
Middleton FA, Strick PL (1996) The temporal lobe is a target of output from the basal ganglia. Proc Natl Acad Sci USA 93(16):8683–8687
Miller JH, Page SE (2009) Complex adaptive systems: an introduction to computational models of social life. Princeton University Press, Princeton
Newell A (1990) Unified theories of cognition. Harvard University Press, Cambridge
Nowak MA, Komarova NL, Niyogi P (2001) Evolution of universal grammar. Science 291(5501):114–118
Orr MG, Lebiere C, Stocco A, Pirolli P, Pires B, Kennedy WG (2018) Multi-scale resolution of cognitive architectures: A paradigm for simulating minds and society. International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. Springer, New York, pp 3–15
Pais D, Caicedo-Nunez CH, Leonard NE (2012) Hopf bifurcations and limit cycles in evolutionary network dynamics. SIAM J Appl Dyn Syst 11(4):1754–1784
Pires B, Crooks AT (2017) Modeling the emergence of riots: a geosimulation approach. Comput Environ Urb Syst 61:66–80
Polinsky M, Van Everbroeck E (2003) Development of gender classifications: modeling the historical change from latin to french. Language, pp 356–390
Prietula M, Carley K, Gasser L (1998) Simulating organizations: computational models of institutions and groups, vol 1. The MIT Press, Cambridge
Rao AS, Georgeff MP (1995) Bdi agents: from theory to practice. ICMAS 95:312–319
Reitter D, Lebiere C (2010) Accountable modeling in act-up, a scalable, rapid-prototyping ACT-R implementation. In: Proceedings of the 2010 international conference on cognitive modeling
Reitter D, Lebiere C (2012) Social cognition: Memory decay and adaptive information filtering for robust information maintenance. In: Proceedings of the twenty-sixth AAAI conference on artificial intelligence, AAAI, pp 242–248

Reynolds CW (1987) Flocks, herds and schools: a distributed behavioral model. ACM SIGGRAPH Comput Gr 21(4):25–34
Ritter F, Haynes SR, Cohen M, Howes A, John B, Best B, Lebiere C, Jones RM, Crossman J, Lewis RL, St Amant R, McBride SP, Urbas L, Leuchter S, Vera A (2012) High-level behavior representation languages revisited. In: Proceedings of the twenty-sixth AAAI conference on artificial intelligence, AAAI, pp 242–248
Romero O, Lebiere C (2014) Simulating network behavioral dynamics by using a multi-agent approach driven by ACT-R cognitive architecture. In: Proceedings of the behavior representation in modeling and simulation conference
Sakellariou I, Kefalas P, Stamatopoulou I (2008) Enhancing netlogo to simulate BDI communicating agents. SETN, Springer, New York, pp 263–275
Schelling TC (1969) Models of segregation. Am Econ Rev 59(2):488–493
Schmidt B (2000) The modelling of human behaviour: the PECS reference models. SCS-Europe BVBA, Delft
Schultz W (2002) Getting formal with dopamine and reward. Neuron 36(2):241–263
Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275(5306):1593–1599
Sierhuis M, Clancey WJ, Van Hoof RJ (2007) Brahms: a multi-agent modelling environment for simulating work processes and practices. Int J Simul Process Model 3(3):134–152
Simon HA (1962) The architecture of complexity. Proc Am Philos Soc 106(6):467–482
Simon HA (1991) Bounded rationality and organizational learning. Organ Sci 2(1):125–134
Stocco A (2018) A biologically plausible action selection system for cognitive architectures: implications of basal ganglia anatomy for learning and decision-making models. Cognit Sci 42:457–490. https://doi.org/10.1111/cogs.12506
Stocco A, Lebiere C, Anderson JR (2010) Conditional routing of information to the cortex: a model of the basal ganglias role in cognitive coordination. Psychol Rev 117(2):541–574
Stocco A, Murray NL, Yamasaki BL, Renno TJ, Nguyen J, Prat CS (2017) Individual differences in the simon effect are underpinned by differences in the competitive dynamics in the basal ganglia: an experimental verification and a computational model. Cognition 164:31–45
Sun R (2006) Cognition and multi-agent interaction: from cognitive modeling to social simulation. Cambridge University Press, Cambridge
Sutton RS (1988) Learning to predict by the methods of temporal differences. Mach Learn 3(1):9–44
Vallacher RR, Read SJ, Nowak A (2017) Computational social psychology. Routledge, London
West R, Nagy N, Karimi F, Dudzik K (2017) Detecting macro cognitive influences in micro cognition: Using micro strategies to evaluate the sgoms macro architecture as implemented in ACT-R. In: Proceedings of the 15th international conference on cognitive modeling, pp 235–236
West RL, Lebiere C (2001) Simple games as dynamic, coupled systems: randomness and other emergent properties. Cognit Syst Res 1(4):221–239
West RL, Stewart TC, Lebiere C, Chandrasekharan S (2005) Stochastic resonance in human cognition: ACT-R vs. game theory, associative neural networks, recursive neural networks, q-learning, and humans. In: Proceedings of the 27th annual conference of the cognitive science society, Lawrence Erlbaum Associates, Mahwah, pp 2353–2358
Wilson EO (2000) Sociobiology. Harvard University Press, Cambridge

Mark G. Orr Research Associate Professor, Biocomplexity Institute & Initiative, University of Virginia. Originally trained as an experimental cognitive psychologist (Univ. of Illinois, Chicago), Dr. Orr was trained in computational psychology (Carnegie Mellon), behavioral neuroscience (Albert Einstein) and epidemiology-agent-based methods (Columbia University). Over the past 8 years, his driving research agenda focused on integrating psychological first principles into agent-based modeling approaches for HHS and DoD. For the past 4 years, he has worked in a high-throughput laboratory that is on the cutting-edge of the agent-based modeling world because it leverages very technical fields, (supercomputing, computer science, network science, physics) in addition to substantive fields (neuroscience, psychology, sociology, epidemiology).

Christian Lebiere is Research Faculty in the Psychology Department at Carnegie Mellon University. He received his B.S. in Computer Science from the University of Liege (Belgium) and his M.S. and Ph.D.

from the School of Computer Science at CMU. During his graduate career, he studied connectionist models and was co-developer of the Cascade-Correlation learning algorithm. Since 1991, he has developed the ACT-R cognitive architecture and co-authored with John Anderson the book The Atomic Components of Thought. His research interests include cognitive architectures and their applications to cognitive psychology, artificial intelligence, human-computer interaction, decision-making, intelligent agents, cognitive robotics and network science.

Andrea Stocco is an Assistant Professor at the Department of Psychology and the Institute for Learning and Brain Sciences (I-LABS) at the University of Washington, Seattle. Since graduating in 2005, he has been working on understanding the human cognitive architecture, using a combination of modeling approaches (both neural networks and ACT-R models), behavioral data, and neuroimaging techniques (fMRI and EEG). In 2013, he also demonstrated the first brain-to-brain interface in humans, using TMS and EEG. His research interests include cognitive architectures, reinforcement learning, and computational approaches to understanding mental disease.

Peter Pirolli is a Senior Research Scientist at the Institute for Human and Machine Cognition. His research involves a mix of cognitive science, artificial intelligence, and human-computer interaction, with applications in digital health, sensemaking, information foraging and online social behavior, among other things. He has a doctorate in cognitive psychology from Carnegie Mellon University and is a Fellow of the National Academy of Inventors, the American Association for the Advancement of Science, the American Psychological Association (Div 3 and Div 21), the Association for Psychological Science, the National Academy of Education, and the ACM Computer-Human Interaction Academy.

Bianica Pires Ph.D., is a Research Assistant Professor at the Biocomplexity Institute of Virginia Tech. She received her PhD in Computational Social Science from George Mason University. Her research involves understanding and exploring complex social systems using various computational methods, including agent-based modeling, social network analysis, and geographic information systems. She has developed models to simulate diffusion and social influence processes at the individual-level that integrate social and cognitive theories of human behavior with empirical data on the local dynamics and movement of agents over a real world physical environment.

William G. Kennedy is an Associate Professor and the Senior Science Advisor in the Center for Social Complexity at George Mason University in Fairfax, Virginia. He holds a B.S. in mathematics from the U.S. Naval Academy, and M.S. in Computer Science from the Naval PostGraduate School, and a Ph.D. in information technology (artificial intelligence) from George Mason University (2003). He is also a retired Navy Captain and a retired federal civil servant. Dr. Kennedy currently conducts research in the field of cognitive science spanning individuals to societies.

Affiliations

Mark G. Orr[1] · Christian Lebiere[2] · Andrea Stocco[3] · Peter Pirolli[4] · Bianica Pires[5] · William G. Kennedy[6]

[1] Biocomplexity Institute & Iniative, University of Virginia, 995 Research Park, Charlottesville, VA 22911, USA

[2] Carnegie Mellon University, Pittsburgh, USA

[3] University of Washington, Seattle, USA

[4] Institute for Human and Machine Cognition, Pensacola, USA

[5] Virginia Tech, Arlington, USA

[6] George Mason University, Fairfax, USA