Journal of Mathematical Psychology 88 (2019) 48–57

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Bidirectional constraint satisfaction in rational strategic decision making

Sudeep Bhatia [a,*,1], Russell Golman [b]

[a] Department of Psychology, University of Pennsylvania, Philadelphia, USA
[b] Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, USA

HIGHLIGHTS

- Bidirectional associative memory is used to study game theoretic decision making.
- The network makes decisions through constraint satisfaction.
- Long run activation is restricted to the set of rationalizable strategies.
- Any pure strategy profile in a stable state must be a pure strategy Nash equilibrium.

ARTICLE INFO

Article history:
Received 13 March 2018
Received in revised form 6 September 2018
Available online xxxx

Keywords:
Decision making
Constraint satisfaction
Neural networks
Game theory
Nash equilibrium

ABSTRACT

We describe the properties of a constraint satisfaction network that is able to reason and decide rationally in strategic games. We use the structure of Bidirectional Associative Memory (BAM), a minimal two-layer recurrent neural network, and assume that network layers represent *self* and *other* strategies, whereas connection weights encode best responses. We apply BAM to finite-strategy two-player games, and show that network activation in the long run is restricted to the set of rationalizable strategies. The network is not guaranteed to reach a stable activation state, but any pure strategy profile that constitutes a stable state in the network must be a pure strategy Nash equilibrium. We illustrate the properties of the network using the traveler's dilemma, the rock–paper–scissors game, and coordination games. The network's behavior also depends on starting activation states, and we show how biases in these starting states can resolve equilibrium selection problems. Strategic decision making is a key part of complex social behavior, and our results illustrate how bidirectional constraint satisfaction networks can perform rational computations in this domain.

1. Introduction

Rationality plays an important role in the study of high-level cognition. Many theories of reasoning, thinking, judgment and decision making rely critically on the notion of a rational agent who is able to encode and process information in an optimal manner. In some settings, the decision maker is assumed to be a rational agent, and behavioral predictions are obtained by analyzing the behavior of this agent (Anderson, 1990; Oaksford & Chater, 2007; Tenenbaum, Griffiths, & Kemp, 2006). In other settings, the behavior of the rational agent is juxtaposed with human behavior, and the differences between human behavior and rational behavior are of theoretical interest (Kahneman & Tversky, 1979; Tversky & Kahneman, 1974). In either case, a complete account of behavior in a given domain requires an understanding of how rational thought and action in that domain could be generated by psychologically plausible cognitive mechanisms.

One area of high-level cognition in which rational behavior is well understood involves game theoretic strategic decision making. The standard decision task studied by game theorists involves two or more players, each allowed to choose between two or more decision strategies. Importantly, the rewards for the players do not depend only on their own choices, but also on the choices of the other players, so that good decision making involves reasoning through reward contingencies, predicting what the opposing players will choose, and best responding to their expected choices. Common knowledge of rationality implies that players will choose only *rationalizable strategies*, which are by definition best responses to other rationalizable strategies (Bernheim, 1984; Pearce, 1984). The Nash equilibrium, a set of (rationalizable) strategies chosen by players so that individual players are unable to benefit by unilaterally changing their strategies, is commonly accepted as

* Corresponding author.
E-mail addresses: bhatiasu@sas.upenn.edu (S. Bhatia), rgolman@andrew.cmu.edu (R. Golman).

1 Funding for Sudeep Bhatia was received from the National Science Foundation grant SES-1626825.

https://doi.org/10.1016/j.jmp.2018.11.004
0022-2496/

the fundamental game theoretic solution concept incorporating rationality and strategic sophistication. Nash equilibrium has been applied to the study of economies, political systems, and societies, and is a fundamental theoretical construct in social science (Camerer, 2003; Colman, 2003; Hart, 1992; Luce & Raiffa, 1957).

Despite its importance, it is not clear how Nash behavior (or rationalizable behavior more generally) could be generated by psychologically plausible cognitive mechanisms, such as those commonly used to describe cognition and behavior in non-strategic domains. The goal of this paper is to address this problem using a constraint satisfaction neural network modeling framework. Our approach is based on Kosko's (1988) Bidirectional Associative Memory (BAM) network, a minimal two-layer recurrent neural network that has traditionally been applied to pattern matching and classification problems in computer science and engineering (see Howard & Kahana, 2002, for a related application to human memory). BAM has binary activation functions, and activation in the network spreads sequentially between the two layers. Processing in the network terminates once activation stabilizes, with final activation states determining the network's responses. In our analysis, we assume that the two layers in the BAM network represent the strategies available to the two players (*self* and *other*), with each node corresponding to a single strategy for a given player. The connections between the nodes in the two layers encode best responses, so that a node in *self*'s layer sends activation to a node in the *other*'s layer only if the strategy represented by the second node is a best response to the strategy represented by the first node.

The BAM network is an analytically tractable generalization of the Hopfield network (Hopfield, 1982), and can perfectly mimic the Hopfield network when both layers have the same number of nodes and node updating across the two layers is simultaneous. As in the Hopfield network, decision making in BAM occurs through constraint satisfaction. Constraint satisfaction networks are well-known class of computational models, with a long history in cognitive science research (see e.g. McClelland, Botvinick, Noelle, Plaut, Rogers, Seidenberg, & Smith, 2010; Read, Vanman, & Miller, 1997 for an overview). They are commonly used to model cognition in non-strategic settings, including perception, causal reasoning, stereotype formation, analogical mapping, legal reasoning, person construal, and preferential choice (Bhatia, 2016; Glöckner & Betsch, 2008; Glöckner, Hilbig, & Jekel, 2014; Holyoak & Powell, 2016; Holyoak & Simon, 1999; Holyoak & Thagard, 1989; Kunda & Thagard, 1996; McClelland & Rumelhart, 1981; Mischel & Shoda, 1995; Simon, Krawczyk & Holyoak, 2004; Simon, Snow & Read, 2004; Thagard, 1989). Many of these constraint satisfaction models also adopt a bidirectional structure. For example, the constraint satisfaction networks proposed by Glöckner and Betsch (2008) and Holyoak and Simon (1999) assume that decision makers have two primary layers, with the first layer encoding cues and the second layer encoding responses. Cues activate responses, but responses can also activate cues, generating a range of anomalous behaviors in judgment and decision making tasks. The BAM network has also been applied to these tasks, and has been shown to mimic preexisting models, while facilitating an analytically grounded understanding of their properties, and the implications of these properties for human behavior (Bhatia, 2016).

Most prior work applying constraint satisfaction networks like BAM to the study of decision making implicitly assumes that long-term activation in the network is necessarily described by a stable state. We find that our implementation of game theoretic decision making in the BAM network restricts long-term activation to the set of rationalizable strategies. However, as best response connections are not necessarily symmetric, there are settings in which the network never achieves a stable activation state. In settings in which the network stabilizes with the activation of a particular strategy profile, this profile is guaranteed to be some pure strategy Nash equilibrium. Conversely, every pure strategy Nash equilibrium corresponds to some stable state in the BAM network. In other words, if a pure strategy Nash equilibrium exists in a given game, the BAM network is able (though not guaranteed) to select it, and likewise if BAM network does select a strategy, it is guaranteed to be in a Nash equilibrium. Even if the BAM network does not select a single strategy, the strategies it oscillates between are all rationalizable.

We illustrate these properties of the BAM network with three examples: The travel's dilemma, the rock–papers–scissors game, and coordination games. The first of these games consists of only a single rationalizable strategy for each player, which corresponds to a pure strategy Nash equilibrium, and we show that the BAM network is guaranteed to stabilize with the activation of nodes in this equilibrium. The second game does not involve any pure strategy Nash equilibrium, but every strategy is rationalizable. We show that the BAM network will display an oscillating pattern of activation in this game, in which each strategy node activates consecutively. Finally, the third game involves multiple pure strategy Nash equilibria. We show that each of these equilibria correspond to a stable state in our network, and that the network is guaranteed to stabilize with the activation of nodes corresponding to one of these equilibria.

We also examine how equilibrium selection problems in coordination games with payoff dominating equilibria can be resolved by the BAM network (Colman, 2003; Harsanyi & Selten, 1988). Payoff dominating equilibria are unambiguously better for all decision makers. With appropriately biased starting states (reflecting attentional biases in favor of desirable strategies), we show that the BAM network necessarily stabilizes with the activation of payoff dominating strategies. Related activation biases also help specify the role of strategy prominence in coordination, and can help generate (arguably rational) choices favoring prominent strategies in coordination games with payoff symmetric equilibria (see Schelling's 1960 argument for focal points).

Finally, we outline extensions of the BAM network that would allow us to make precise (probabilistic) predictions regarding behavior. People may occasionally violate these predictions, as Nash equilibrium itself is not a great behavioral model (Camerer, 2003). We thus also suggest modifications to the BAM network that may allow it to generate behaviorally plausible (though not necessarily rational) choices.

Overall, our results show how constraint satisfaction approaches to modeling decision making can be used to describe the cognitive basis of rational strategic choice. In doing so, they illustrate the theoretical power of models of constraint satisfaction, and highlight their applicability to the study of complex social behavior.

2. Game theoretic decision making

In strategic games, two or more players make choices over a set of strategies. Crucially, the strategies chosen by the players collectively determine the outcomes of the game, so that each player's utility depends on the *other*'s choice as well as on their own. We define a finite-strategy two-player game with a set of pure strategies for each player, $S_1 = \{s_{11}, \dots, s_{1N}\}$ and $S_2 = \{s_{21}, \dots, s_{2M}\}$ respectively, and a pair of payoff functions u_1 and u_2 that give each player's utility for each profile of pure strategies (s_{1i}, s_{2j}) (see, e.g., Hart, 1992). Thus if player 1 selects s_{1i} and player 2 selects$s_{2j}$the utility for player 1 is $u_1(s_{1i}; s_{2j})$ and the utility for player 2 is $u_2(s_{2j}; s_{1i})$. We use the notation $\boldsymbol{u}_{ij}$ as a shortcut for $\left(u_1(s_{1i}; s_{2j}), u_2(s_{2j}; s_{1i})\right)$.

The most standard solution concept for a strategic game is Nash equilibrium, which relies on common knowledge of rationality and accurate expectations. A Nash equilibrium is a strategy profile in which no player can obtain higher utility by unilaterally changing his strategy; each player is already playing a best response to the equilibrium strategy profile. We define the set of best responses for player μ to an opponent's strategy $s_{-\mu}$ as $\text{BR}\left(s_{-\mu}\right) = \arg\max u_\mu(s_\mu; s_{-\mu})$. Then a pure strategy Nash equilibrium can be defined as a strategy profile (s_{1i}, s_{2j}) such that $s_{1i} \in \text{BR}(s_{2j})$ and $s_{2j} \in \text{BR}(s_{1i})$.

The concept of Nash equilibrium can be generalized to relax the assumption that players somehow have correct expectations about what others will do. The solution concept of rationalizability (Bernheim, 1984; Pearce, 1984) retains the assumption of common knowledge of rationality, but imposes no additional constraints on behavior. As in a Nash equilibrium, players best respond to the strategy they expect their opponent to select, but in contrast to a Nash equilibrium, this expectation is not necessarily correct. Players must only be able to "rationalize" their strategy choice as a best response to one of the opponent's rationalizable strategies. We define the set of rationalizable strategies for each player as the maximal sets R_1 and R_2 such that any $s_1 \in R_1$ satisfies $s_1 \in \text{BR}(s_{2j})$ for some $s_{2j} \in R_2$ and any $s_2 \in R_2$ satisfies $s_2 \in \text{BR}(s_{1i})$ for some $s_{1i} \in R_1$. Clearly, any Nash equilibrium profile is rationalizable, and if the sets of rationalizable strategies are singletons, then these strategies form a Nash equilibrium.

3. Bidirectional associative memory

Nash equilibrium and rationalizability are two of the most important solution concepts in game theory. Here we examine how choices corresponding to these concepts can be generated by rational decision makers, modeled with constraint satisfaction neural networks. These types of networks rely on bidirectional (recurrent) connectivity between their component nodes, which is able to generate sophisticated dynamics and subsequently explain a range of human behavior. Although these networks have traditionally been used only to explain behavior in non-strategic settings (as outlined in the introduction above) they can be applied to strategic game theoretic decision making. Indeed, these networks are particularly suitable for this task, as game theoretic decision making features complex interactions between the choices of different decision makers; interactions that can be specified using recurrent bidirectional connectivity.

The Bidirectional Associative Memory Network (BAM) is a particularly powerful (and mathematically tractable) constraint satisfaction neural network (Kosko, 1987, 1988). It consists of two layers with binary connections between their respective nodes and binary activation functions for any given node. When the connections between its nodes are symmetric then BAM is guaranteed to reach a stable pattern of activation, regardless of its starting state. This property has been used by scholars to solve a variety of practical tasks involving associative memory and pattern completion (Kosko, 1988) and also model non-strategic human decision making, and the biases that it often involves (Bhatia, 2016).

In this paper, we model how an individual decision maker can reason through two-player finite strategy games, using the BAM network. Strategies in these games for each of the two players can be represented in each of BAM's two layers. We will assume that the first layer in the BAM network represents strategies available to player 1 (or *self*). If player 1 can choose from the set of strategies $S_1 = \{s_{11}, \ldots, s_{1N}\}$, then the first layer in our network consists of N nodes, with node i representing strategy s_{1i}. The activated nodes in this layer represent the strategies that the decision maker considers playing in the game.

Correspondingly we will assume that the second layer in the BAM network represents strategies available to player 2 (the opponent, or *other*). If player 2 can choose from the set of strategies $S_2 = \{s_{21}, \ldots, s_{2M}\}$, then the second layer in our network consists of M nodes, with node j representing strategy s_{2j}. The activated nodes in this layer represent the strategies that the decision maker thinks *other* might play in the game.

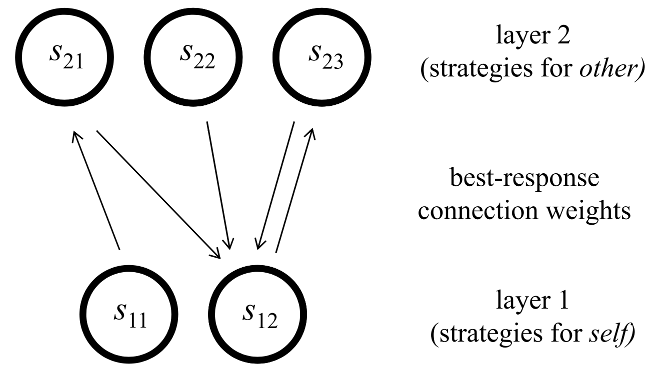

Fig. 1. Example of a BAM network. Here node s_{1i} corresponds to strategy i for *self* and node s_{2j} corresponds to strategy j for *other*. Connection weights encode best responses. Here we have a game with two strategies for *self* and three strategies for *other*, with strategy 2 for *self* serving as a best response to strategies 1, 2 and 3 for *other*, strategy 1 for *other* serving as a best response to strategy 1 for *self*, and strategy 3 for *other* serving as a best response to strategy 2 for self. Note that the only Nash equilibrium in the game encoded in this network is (s_{12}, s_{23}).

As mentioned earlier, node activation in BAM is binary, with each node being on or off. We will assume that every node has the same binary activation function, with activation triggered by strictly positive input. For any node k (in either layer of the network) with input I_k, the activation function f_k is specified by Eq. (1).

$$f_k(I_k) = \begin{cases} 1 \; if \; I_k > 0; \\ 0 \; if \; I_k \leq 0. \end{cases} \tag{1}$$

In a slight abuse of notation, we denote by f the activation function for either layer of the network, with components f_k for every node k.

Connections between nodes are also binary, with each node in the first layer either connected or not connected to each node in the second layer, and each node in the second layer either connected or not connected to each node in the first layer. There are no connections between two nodes in one layer. The network structure can thus be described by the matrices $\mathbf{W^{12}}$ and $\mathbf{W^{21}}$ which represent connections from layer 1 to layer 2, and from layer 2 to layer 1 respectively.

We will assume the pattern of connections in our network captures best responses. Particularly, if $s_{2j} \in \text{BR}(s_{1i})$ then we assume the connection from node i in the first layer to node j in the second layer is $W_{ij}^{12} = 1$. If $s_{2j} \notin \text{BR}(s_{1i})$ then we assume $W_{ij}^{12} = 0$.We assume a similar pattern of connectivity from the second layer to the first, so that $W_{ji}^{21} = 1$ if $s_{1i} \in \text{BR}(s_{2j})$ and $W_{ji}^{21} = 0$ otherwise. In essence, an activated strategy s_{1i} for *self* sends positive inputs to strategies s_{2j} for *other* that serve as best responses to s_{1i}. Conversely, an activated strategy s_{2i} for *other* sends positive inputs to strategies s_{1i} for *self* that serve as best responses to s_{2j}. Fig. 1 provides an illustration of the proposed network.

We write the activation of any node i in the first layer, at time t, as $A_{1i}(t)$, and any node j in the second layer, at time t, as $A_{2j}(t)$. With the connectivity specified above, vectors $\boldsymbol{A_1}(t)$ and $\boldsymbol{A_2}(t)$ together represent network activation at time t. We can describe their dynamics using Eqs. (2a) and (2b).

$$\boldsymbol{A_1}(t) = f(\boldsymbol{A_2}(t-1)\cdot\mathbf{W^{21}}) \tag{2a}$$

$$\boldsymbol{A_2}(t) = f(\boldsymbol{A_1}(t)\cdot\mathbf{W^{12}}) \tag{2b}$$

As formalized in the above equation, node updating in our network is sequential, with layer 1 updating before layer 2. This does not affect the network's behavior, except at the starting point $t = 0$. Here, the above assumption implies that the network begins processing the decision when some nodes in layer 1 are activated exogenously (intuitively, *self* begins the decision process by first considering his strategies). The activation of strategies in layer 1 then activates the strategies in layer 2 which then alters the pattern of activation in layer 1. In essence, deliberation involves iterative activation of strategies that serve as best responses to previously activated strategies. Fig. 2 illustrates the spread of activation in the BAM network from Fig. 1. The choice of the starting point activation in our network can affect subsequent node activation and sometimes will determine selection among multiple stable states.

Like related recurrent neural network models, the BAM network can make some decisions through constraint satisfaction, that is, by settling into a stable activation state. A stable activation state in the network is a state from which endogenous deviations are not possible. Activation states $\boldsymbol{A}_1^*$ in layer 1 and $\boldsymbol{A}_2^*$ in layer 2 are stable if $\boldsymbol{A}_1^*(t) = \boldsymbol{A}_1^*(t+1)$ and $\boldsymbol{A}_2^*(t) = \boldsymbol{A}_2^*(t+1)$. We assume that a decision maker chooses one of the strategies that are activated in layer 1 and expects *other* to choose one of the strategies activated in layer 2, in the network's stable state.

The connections assumed in this paper are not necessarily symmetric, as strategy s_{1i} can be a best response to strategy s_{2j} without s_{2j} being a best response to strategy s_{1i}. This means that the network is not always guaranteed to stabilize. If the network does not stabilize, then it enters a pattern of oscillating activation in which a certain subset of nodes are activated and deactivated consecutively. We assume that the nodes that are activated (but then deactivated) as part of this oscillating pattern correspond to the set of strategies from which decision makers make their final choice. Nodes that are not activated as part of this oscillating pattern correspond to strategies that are ignored by the decision maker.

Long-run activation of the BAM network does not always uniquely determine a choice prediction. It does rule out non-rationalizable strategies, which is the model's primary testable prediction. Still, we might like a fully specified choice rule. Later in this paper, we consider various extensions that allow the model to make more precise behavioral predictions.

4. Stability and Nash equilibrium

First, suppose the network reaches a stable state corresponding to a pure strategy profile, that is, a stable state of activation in which only one node is activated in each layer of the network. Our first result characterizes that stable activation state as corresponding to a Nash equilibrium.

Proposition 1. *Suppose that* $\lim_{t\to\infty} (\boldsymbol{A}_1(t), \boldsymbol{A}_2(t)) = (\boldsymbol{A}_1^*, \boldsymbol{A}_2^*)$ *with unique nodes i and j in each layer for which* $A_{1i}^* > 0$ *and* $A_{2j}^* > 0$. *Then* (s_{1i}, s_{2j}) *is a Nash equilibrium.*

Proof. Strategy s_{1i} is activated at time t (i.e., $A_{1i}(t) > 0$) if and only if there exists s_{2j} such that $s_{1i} \in \mathrm{BR}(s_{2j})$ and $A_{2j}(t-1) > 0$, and conversely, strategy s_{2j} is activated at time t (i.e., $A_{2j}(t) > 0$) if and only if there exists s_{1i} such that $s_{2j} \in \mathrm{BR}(s_{1i})$ and $A_{1i}(t-1) > 0$. Thus, if the network converges to $(\boldsymbol{A}_1^*, \boldsymbol{A}_2^*)$ with unique nodes i and j in each layer for which $A_{1i}^* > 0$ and $A_{2j}^* > 0$, then $s_{1i} \in \mathrm{BR}(s_{2j})$ and $s_{2j} \in \mathrm{BR}(s_{1i})$, which means that (s_{1i}, s_{2j}) is a Nash equilibrium. □

Proposition 1 tells us that the neural network will find a Nash equilibrium if it is able to converge on a single strategy profile. The process through which the network finds this Nash equilibrium is constraint satisfaction. The example of the *traveler's dilemma* in the section below illustrates how the network converges on a pure strategy Nash equilibrium.

The game of *rock–paper–scissors* discussed in the section below illustrates that the network may not converge to a stable state. Even if the network does not reach a stable state, however, we can characterize the nodes which may experience recurrent activation. In the long run, activation is restricted to the set of rationalizable strategies, R_1 and R_2 respectively. Let $\overline{A}_1(t)$ and $\overline{A}_2(t)$ respectively denote the sets of strategies that are activated at time t, i.e., $s_{1i} \in \overline{A}_1(t)$ if $A_i^1(t) > 0$.

Theorem 1. *There exists* τ *such that for any* $t > \tau$ *we have* $\overline{A}_1(t)\, R_1$ *and* $\overline{A}_2(t) \subseteq R_2$.

Proof. We show that if $s_{1i} \notin R_1$, then for large enough t, $A_{1i}(t) = 0$. (The argument for player 2's strategies is analogous.) If $s_{1i} \notin R_1$, then any chain of best responses can include s_{1i} at most once. Strategy s_{1i} is activated at time t (i.e., $A_{1i}(t) > 0$) if and only if there exists $s_{2j} \notin R_2$ such that $s_{1i} \in \mathrm{BR}(s_{2j})$ and $A_{2j}(t-1) > 0$. The players only have a finite number $(N+M)$ of strategies, so for $t > N+M$, there are no more strategies available to seed a chain of best responses, so $A_{1i}(t) = 0$. □

Theorem 1 tells us that the neural network will select only rationalizable strategies. Thus, strategic rationality emerges from the structure of the bidirectional network. The network may never converge to a state with stable activation, so we may not be able to identify a single strategy that will necessarily be chosen, but we can make testable predictions about what will *not* be chosen.

We can also recognize that the lack of a point prediction creates space for contextual factors to matter. An individual's eventual decision may depend on which strategy he considers first, which could in turn depend on the salience of different strategies, how the strategies are framed, or how the decision maker's attention is anchored. The *coordination game* discussed in the section below illustrates how the starting point determines which of the multiple Nash equilibria is eventually selected by the network.

It is straightforward to observe that any pure strategy Nash equilibrium would constitute a stable state in our network. The strategies in the Nash equilibrium would, due to the nature of the connection weights, reinforce each other and, once activated, sustain their activation.

Theorem 2. *If* (s_{1i}, s_{2j}) *is a Nash equilibrium, then there exists a stable state* $(\boldsymbol{A}_1^*, \boldsymbol{A}_2^*)$ *with unique nodes i and j in each layer for which* $A_{1i}^* > 0$ *and* $A_{2j}^* > 0$.

Proof. This follows from our assumption that strategies are connected to their best responses. □

Theorem 2 tells us that a Nash equilibrium is in fact a dynamic equilibrium, a stable stationary state of the network. If an equilibrium is made salient to a decision maker behaving in accordance with the BAM network, the individual will indeed choose his equilibrium strategy.

5. Illustrations

In this section we apply the BAM network to three representative games. The games vary in the number of Nash equilibria and in the size of the set of rationalizable strategies. These examples demonstrate that activation in the BAM network in different contexts either may converge to a unique pure strategy Nash equilibrium from any initial state, may fail to converge at all as it oscillates through multiple rationalizable strategies, or may converge to one of many stable profiles depending on the initial state.

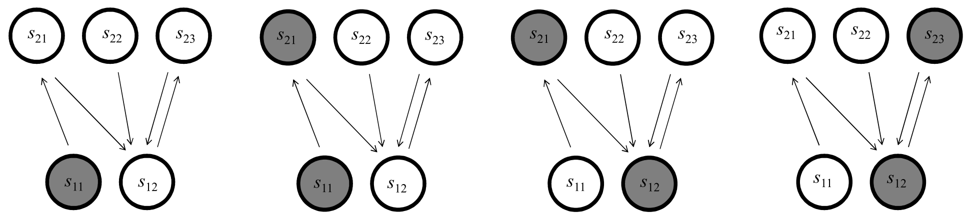

Fig. 2. Example of spread of activation in the BAM network shown in Fig. 1. Here the decision begins with the activation of strategy 1 for *self*. This activates strategy 1 for *other*, which in turn activates strategy 2 for *self*. Finally strategy 2 for *self* activates strategy 3 for *other*. This is a stable activation state (and corresponds to the only Nash equilibrium in the encoded game).

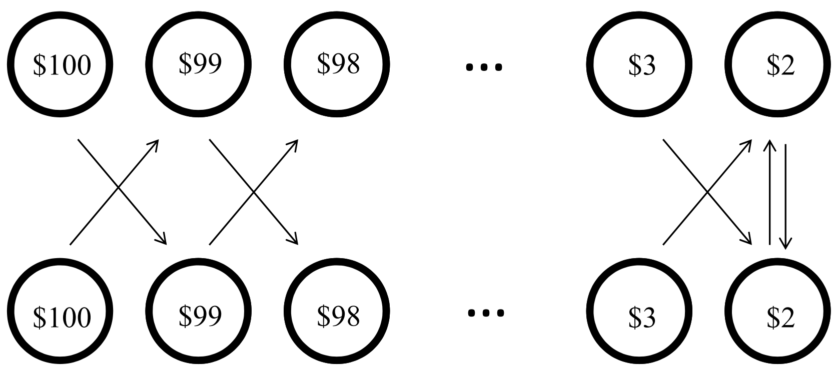

Fig. 3. Example of a BAM network encoding the traveler's dilemma game. Here the best response for *self* is to claim $1 less than *other*, and vice versa. Node activation in this network is guaranteed to terminate with the activation of the $2 strategy nodes for *self* and *other*.

5.1. Traveler's dilemma

The traveler's dilemma is a generalization of the famous prisoner's dilemma, conceived in order to demonstrate unraveling in a strategic game (Basu, 1994). In the original parable, two travelers have lost identical antiques and must request compensation between $2 and $100. The airline (which is responsible for the lost luggage) will accept the lower claim as valid and pay that amount to both players, and, to deter lying, will penalize the higher claimant with a $2 fee and will reward the lower claimant with $2 bonus. We represent the game with the strategy sets $\boldsymbol{S}_1 = \boldsymbol{S}_2 = \{2, 3, \ldots, 100\}$, where x_{1i} and x_{2j} correspond to the dollar amounts associated with strategies s_{1i} and s_{2j}, and with the following utilities:

$$\boldsymbol{u}_{ij} = \begin{cases} \left(x_{2j} - 2, x_{2j} + 2\right) & \textit{if } x_{1i} > x_{2j}; \\ \left(x_{1i}, x_{2j}\right) & \textit{if } x_{1i} = x_{2j}; \\ \left(x_{1i} + 2, x_{1i} - 2\right) & \textit{if } x_{1i} < x_{2j}. \end{cases} \tag{3}$$

The airline's scheme, of course, does not actually reward honesty; it rewards undercutting the other traveler. The best response is always to claim exactly $1 less than the other traveler does (if it is feasible to do so). As it turns out, the only rationalizable strategy for either player is to claim $2, and the unique Nash equilibrium has both players claim $2.

The network connectivity implied by the traveler's dilemma is illustrated in Fig. 3. Given enough time this network is guaranteed to stabilize with the activation of the node corresponding to $2 in layer 1, the node corresponding to $2 in layer 2, and the deactivation of all the other nodes (corresponding to higher claims). If, for example, the deliberation process begins with *self* considering claiming $100, i.e., node $100 being activated in layer 1, then node $99 will become activated in layer 2, and in turn node $98 will become activated in layer 1, and so on, until only the nodes corresponding to $2 in each layer are activated. The unique Nash equilibrium corresponds to the only stable state of activation here, because it consists of the only rationalizable strategy for each player. Intuitively, when a decision maker is given enough time to reason, our model predicts that he will choose to claim $2 and will expect the other player to do so as well.

5.2. Rock–paper–scissors

The classic game of rock–paper–scissors is the simplest symmetric, zero-sum game with non-transitive winning strategies. Each player has three pure strategies: rock, paper, or scissor. The loop is that rock "defeats" scissors, scissors "defeats" paper, and paper "defeats" rock. If both players play the same strategy, then the game is a tie. We can represent the rock–paper–scissors game with the utilities u_{ij} described in Table 1.

The game of rock–paper–scissors has no pure strategy Nash equilibrium, but every strategy is rationalizable. Every strategy is a best response to some other strategy, but no strategy is a best response to itself.

The network connectivity implied by the rock–paper–scissors game is illustrated in Fig. 4. With a single node initially activated, the network will never stabilize, regardless of which node is the starting point. Instead, the network will display an oscillating pattern of activation, in which each node in each layer activates consecutively. If, for example, the network begins with the activation of the rock node in layer 1, then the node corresponding to paper (the best response to rock) will activate in layer 2. In turn, the node corresponding to scissors will then activate in layer 1. This leads to activating rock in layer 2, and so on. Intuitively, our model predicts that a decision maker will cycle through all three possible strategies as he reasons about the game, and any of these strategies may eventually be chosen.

5.3. Coordination game

A coordination game captures situations in which the players' primary incentives are to behave similarly, as for example in the case that two friends would like to get together at a meeting place and each has to choose where to go. If they both arrive at the same location, then they each obtain a high reward (e.g., they get to enjoy

Table 1
Utilities, $\boldsymbol{u_{ij}}$, in the rock–paper–scissors game. Each player has three pure strategies: rock, paper, or scissor. The loop is that rock "defeats" scissors, scissors "defeats" paper, and paper "defeats" rock. If both players play the same strategy, then the game is a tie.

		Other Rock	Paper	Scissors
	Rock	(0, 0)	(−1, 1)	(1, −1)
Self	Paper	(1, −1)	(0, 0)	(−1, 1)
	Scissors	(−1, 1)	(1, −1)	(0, 0)

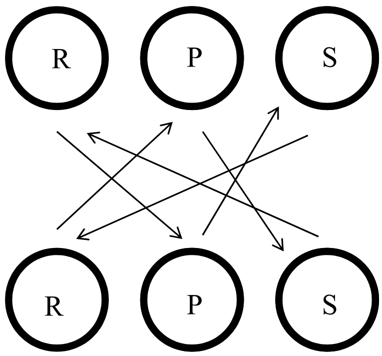

Fig. 4. Example of a BAM network encoding the rock–paper–scissors game. Here rock (R) is a best response to paper (P), which is a best response to scissor (S), which in turn is a best response to rock. Node activation does not stabilize, and instead oscillates with the successive activation of each of these three strategies for *self* and *other*.

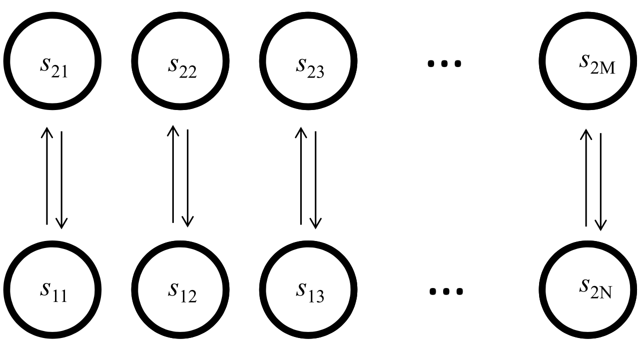

Fig. 5. Example of a BAM network encoding a coordination game. Here the *other's* best response to any strategy played by *self* is the same strategy, and vice versa. The activation of each pair of strategies (s_{1i}, s_{2j}) with $i = j$ is stable in the network.

each other's company). If they arrive at different locations, then they each obtain a low reward or a punishment (e.g., a solitary evening). We can represent a coordination game with $\boldsymbol{S_1} = \boldsymbol{S_2}$ and a utility function that has the following property:

$$\boldsymbol{u}_{ij} = \begin{cases} (v_{1i}, v_{2j}) & \textit{if } s_{1i} = s_{2j}; \\ (0, 0) & \textit{if } s_{1i} \neq s_{2j}. \end{cases} \tag{4}$$

with $v_{1i} > 0$ and $v_{2j} > 0$ for all i and j. In a coordination game, *self* is always incentivized to play the strategy that he expects *other* to play and vice versa, regardless of the specific strategy involved. Thus, there are $N = M$ pure strategy Nash equilibria in the game, with each Nash equilibrium corresponding to an outcome in which *self* and *other* choose the same strategy. As in the rock–paper–scissors game, every strategy is rationalizable. However, whereas the network never stabilizes for the rock–paper–scissors game, it immediately stabilizes for a coordination game.

The network connectivity implied by a coordination game is illustrated in Fig. 5. The network is guaranteed to stabilize with the activation of the same strategy nodes in layers 1 and 2. The precise strategies activated in the stable state depend on the starting point in the deliberation process, so that if deliberation begins with the activation of node i in layer 1 (representing strategy s_{1i} for *self*), then the network will stabilize with the activation of node i in layer 1 and the activation of the corresponding node $j = i$ in layer 2. Any pair of nodes corresponding to a Nash equilibrium creates a mutually reinforcing pattern of activation, a stable state in the network. Intuitively, our model predicts that a decision maker will choose to play the strategy that he first begins thinking about and will expect *other* to play this strategy as well.

6. Salience and starting points

Our analysis of coordination games in the prior section suggests that the starting states BAM network can be used to resolve issues regarding equilibrium selection in such games. Here we explore these issues in more detail in two different types of coordination games.

6.1. Coordination with payoff dominance

Consider first the case of a coordination game with payoff dominance. This game also involves the structure outlined in Eq. (4), however, it further restricts strategy payoffs so that $v_{1i} = v_{2j} \textit{if} s_{1i} = s_{2j}$, that is, each strategy offers both players an identical payoff if they successfully coordinate on that strategy. Moreover, the set of strategy payoffs are structured so that there exists one strategy k^* so that we have $v_{1k^*} = v_{2k^*} > v_{1k} = v_{2k}$ for all $k \neq k^*$. Here, successful coordination on k^* yields higher payoffs to both players than successful coordination on any other strategy, and for this reason, k^* is a payoff dominating Nash equilibrium. In the example with two friends deciding where to go, k^* corresponds to a location that is individually optimal for both friends.

As in the more general coordination game outlined above, each of the $N = M$ strategy pairs in which both players select the same strategy serve as pure strategy Nash equilibria: If the decision maker is a rational utility maximizer he should select k^* only if he expects his opponent to do so as well; but there is no reason to do so, as the opponent faces the same dilemma (see Colman, 2003; Harsanyi & Selten, 1988). While common knowledge of rationality does not restrict the decision maker's choice, we might still think that if all players are rational, they should somehow figure out how to obtain the payoff dominating Nash equilibrium k^*.

The cognitive implementation of game theoretic reasoning in the BAM network is capable of solving the equilibrium selection problem with reasonable assumptions about the psychology of attention. Particularly, if we assume that decision makers are more likely to attend to desirable strategies first, then we can specify starting activation states as being a function of strategy payoffs. For example we can state that strategy i is activated at $t = 0$ if and only if $max_{s_{2j} \in S_2} u_1(s_{1i}; s_{2j}) > max_{s_{2j} \in S_2} u_1(s_{1k}; s_{2j})$ for all $k \neq i$. With such an assumption it is guaranteed that the payoff dominating strategy is activated first, and that the network in turn stabilizes with the activation of this strategy. Attention that is biased in favor of highly rewarding actions has been argued to be a component of rational thought and behavior (see Lieder, Griffiths and Hsu, in press, for an overview). Our analysis shows how such an attentional bias, when implemented within the BAM network, facilitates rational behavior in game theoretic decision making.

6.2. Prominence

Another setting in which the BAM framework can facilitate the selection of optimal strategies involves coordination games with prominent labels. Strategies in these games do not differ in terms of payoffs but rather by exogenous cues that distinguish one strategy label from the rest. An example of such a game involves variants of the game in Eq. (4), with a fixed standard payoff $v_{1i} = v_{2j} = v$ whenever $s_{1i} = s_{2j}$, but with some strategy k^* artificially highlighted or emphasized (Crawford, Gneezy, & Rottenstreich, 2008; Mehta, Starmer, & Sugden, 1994; Schelling, 1960). In the example with two friends deciding where to go, such a strategy could be one that is especially popular or one on which they had coordinated previously.

As with payoff dominance, common knowledge of rationality does not imply choice of the prominent strategy: The decision maker should select k^* only if he expects his opponent to do so as well; but there is no reason to do so, as the opponent faces the same dilemma. Still, we would expect rational players to use strategy prominence to avoid coordination failure. Once again, the BAM framework provides a convenient specification of the effects of prominence in terms of the starting states of the network. It is reasonable to assume that attention is directed towards the prominent strategy, and that this strategy is thus most likely to be activated at the first time period. This leads to the prominent strategy being activated once the network stabilizes. Furthermore, if the opponent also behaves as predicted by the BAM network, then both players would eventually select the prominent strategy and successfully coordinate. This analysis indicates how the BAM implementation of game theoretic reasoning generates rational behavior from core cognitive principles.

7. Model extensions

The above sections present a fairly abstract exposition of the BAM network. We consider only the general properties of the network that give rise to different dynamics in different games. This generality allows us to make mathematically rigorous claims about the relationship between BAM activation and rationalizability, and thus understand how rational strategic decision making can be instantiated within a plausible cognitive model. However, more structure is required in order to make unique behavioral predictions. For example, it is not yet specified what the BAM network chooses when there are multiple possible stable states, when there are multiple strategies activated in a single stable state, or when the network oscillates indefinitely without stabilizing. These settings emerge in games for which rationalizability fails to specify unique choices – i.e. games in which there are multiple Nash equilibria or games in which there are no pure strategy Nash equilibria –, suggesting that in order to make more specific predictions we need to consider extensions of the BAM network that may cause it to go beyond rationalizability.

7.1. Precise probabilistic predictions

We first consider extensions of the BAM network that allow it make precise predictions that facilitate empirical tests and potentially quantitative model fits. Due to the probabilistic nature of choice (see Loomes, 2015, for a review) we consider extensions that give rise to a probability distribution over all possible choices. For a set of N player 1 strategies $S_1 = \{s_{11}, \ldots, s_{1N}\}$ these predictions take the form of an N dimensional vector of probabilities $\boldsymbol{p_1} = [p_{11}, \ldots, p_{1N}]$, such that $\sum_{i=1}^{N} p_{1i} = 1$, with p_{1i} capturing the choice probability of strategy s_{1i}.

The predicted choice probabilities are generated by network's final activation states. These activation states themselves depend on the network's starting point, that is, the strategy that is activated at $t = 0$ (intuitively, the strategy that the decision maker thinks of first). We assume that the network always begins with the activation of a single strategy, that the selection of this strategy is probabilistic, and that activation probabilities depend on the relative salience of the player's strategies. We write the starting point activation probabilities for player 1 as $\boldsymbol{r_1} = [r_{11}, \ldots, r_{1N}]$, with $\sum_{i=1}^{N} r_{1i} = 1$. We interpret r_{1i} as capturing the relative salience of strategy s_{1i} for player 1, and this corresponds to the probability that the network starts with only strategy i activated at $t = 0$ (i.e. $r_{1i} = \text{Prob}[A_{1i}(0) = 1 \text{ and } A_{1k}(0) = 0 \text{ for } k \neq i]$). Relative salience could be a product of payoff magnitude or prominence manipulations, in which case the high-payoff strategy or the prominent strategy would have a higher r_{1i}. Alternatively, it could be the case that there are no salience biases, in which case $r_{1i} = 1/N$ for all i.

Given the specification of starting network activation states we can now derive choice probabilities for the BAM network. We have been thinking of choice as corresponding to long-run activation, but what value of t constitutes the long run? For purposes of making precise predictions, we assume that choice probabilities are determined by the network activation the first time that a network activation state repeats. In the case in which the network stabilizes, this activation state is indeed the long-run stable network activation state. In the case in which the network never stabilizes, this activation state is the first activation state reached in the sequence of activation states that characterizes the network's long-run oscillatory regime. The *first repetition* rule could be implemented with a second set of nodes that track how often each network state becomes activated, and a threshold for triggering choice set at $T = 2$ activations. Actually, because periodic network oscillations repeat deterministically, any value of $T > 1$ will select the same network activation state, and we adopt the first repetition rule ($T = 2$) just for simplicity. We can represent the selected activation state for layer 1 as $\boldsymbol{A_1^R}$. The nodes that are activated in $\boldsymbol{A_1^R}$ correspond to a subset of player 1's rationalizable strategies.

We assume that this activation state combines with a tremble noise term to generate choice probabilities $\boldsymbol{p_1}$, so that the choice probability for player 1's strategy i is given by:

$$p_{1i} = \frac{A_{1i}^R + \theta}{\sum_{k=1}^{N} A_{1k}^R + \theta} \tag{5}$$

where θ is a noise parameter that determines the extent of the tremble noise. If $\theta = 0$, the decision maker always chooses uniformly one of the strategies that are activated in $\boldsymbol{A_1^R}$, whereas if $\theta = \infty$, the decision maker selects each available strategy with an equal probability regardless of the $\boldsymbol{A_1^R}$. For moderate values of θ we obtain higher choice probabilities for strategies activated in $\boldsymbol{A_1^R}$ compared to other strategies (though every strategy has a non-zero probability of being chosen).

When combined with the starting point probabilities and an analysis of BAM dynamics, Eq. (5) can be used to make precise probabilistic predictions for choice in any game that can be represented by the BAM network. Although these predictions depend on the structure of the game, contextual factors that influence starting network activation, and on the degree of tremble noise, we can still generally predict that the strategies that are given the highest choice probabilities are guaranteed to be rationalizable. Thus, our probabilistic extension of the BAM network can still be seen as implementing a rational choice with some noise.

7.2. Predicting Behavior

Of course it is well known that human decision makers display systematic deviations from Nash equilibrium predictions, and often select non-rationalizable strategies with a high probability. These behaviors are thus outside of the explanatory scope of the BAM network. For example, the BAM network involves an iterative reasoning process according to which strategies successively activate their best responses. We assume that this type of reasoning continues until the network stabilizes into either a consistent activation state or else an oscillating activation state (which necessarily includes only rationalizable strategies). Of course, real people do not exhibit so much strategic sophistication. In the aforementioned traveler's dilemma, for example, people often make higher claims, which cannot be supported by any rationalizable strategy (Capra, Holt, Goeree, & Gomez, 1999; Goeree & Holt, 2001). Actual behavior is often better modeled by level-*k* thinking (or cognitive hierarchies), which captures bounded rationality (Camerer, 2003; Nagel, 1995; Stahl & Wilson, 1994). The level-*k* model assumes that players can reason through *k* steps of best response analysis, where the value of *k* may be heterogeneous in a population. Our proposed BAM network could capture this sort of level-*k* thinking if we did not associate decisions with long-run patterns of activation, but rather terminated the network dynamics at a finite time horizon and associated decisions with activation when this horizon was reached. Such an extension to the BAM network would also predict that the extent of strategic sophistication depends on the amount of time that the network is allowed to deliberate, and thus predict that the network's decisions can vary with deliberation time and time pressure.

Yet another deviation from Nash theory involves sensitivity to the payoffs in the game. For example, Capra et al. (1999) find that final choices in the traveler's dilemma are higher with higher values of the fee for the higher claimant and the bonus for the lower claimant, even though the precise value of the fee and bonus does not influence the structure of best responses (the Nash equilibrium is to always claim the lowest amount, regardless of the size of the fee and bonus). Other games, such as the minimum-effort coordination game, the stag hunt game, and the matching pennies game also display this behavioral pattern, so that changing payoffs without altering the structure of best responses (i.e. changing how bad it is to play a suboptimal strategy, but not changing when a strategy is suboptimal) can alter observed choices (Goeree & Holt, 2001). The BAM network cannot capture this type of payoff sensitivity as network connection weights (and thus network activation states, dynamics, and final responses) depend entirely on best responses. Actual behavior in these settings is often better modeled by quantal response equilibrium, which proposes that decision makers may select suboptimal responses to the opponent's expected play, but that the frequency of these mistakes varies inversely with their cost (McKelvey & Palfrey, 1995). This property of quantal response equilibrium introduces a responsiveness to payoff size, and could be incorporated into the BAM network by allowing for continuous connection weights and activation states. For example, connection weights from strategy s_{2j} to s_{1i} could depend not on whether or not s_{1i} is a best response to s_{2j}, but rather on the payoff that would be obtained by playing s_{1i}if the opponent played s_{2j}. This would ensure that changes to the payoffs would be reflected in changes to network connectivity, even if there are no corresponding changes to the best response structure. Likewise, activation states that are continuous, rather than binary, would ensure that continuous changes to network memory influence the extent to which different strategies are activated, and thus the likelihood of playing these strategies. With both payoff-dependent continuous connection weights and continuous activation states, increasing the payoff of a particular strategy would increase the likelihood of choosing that strategy.

In this paper we have avoided complex assumptions regarding finite time horizons for network dynamics or continuous connection weights and activation functions. This has allowed us to make mathematically rigorous claims regarding the relationship between the BAM network and rational behavior. However, in follow up work we are considering these extensions to the BAM network. We find that a pair of stochastic accumulator networks with bidirectional links (what we call the *dual accumulator* model) provides a more accurate characterization of human behavior in game theoretic decision making than the present BAM network, and that it can even outperform sophisticated behavioral theories such as level-*k* reasoning and quantal response equilibrium (Golman & Bhatia, 2017; Golman, Bhatia, & Kane, 2018). These results show that both rational and irrational aspects of game theoretic deliberation can be captured by constrained satisfaction processes in a two-layer recurrent neural network.

8. Discussion

Constraint satisfaction is a key feature of high-level cognition, and models of constraint satisfaction – often formalized using recurrent neural networks – are frequently used to study human reasoning, judgment, and decision making. In this paper we extend this research to game theoretic decision making. In particular, we adapt the Bidirectional Associative Memory (BAM) model (Kosko, 1988), a minimal two-layer recurrent neural network, to make decisions in finite strategy two-player games.

BAM is well-suited for this task. Choices in game theoretic settings are interdependent, with the payoff generated by choosing any one strategy being a function of the choice made by the other decision maker. The recurrent connectivity in BAM can be used to model this type of interdependence. We assume that the two layers in the BAM represent strategies available to the *self* and strategies available to the *other*, and connections between these two layers capture the best responses to the various strategies. With this structure, we show that activation in the BAM network in the long run can only be sustained for rationalizable strategies. Decision making with the BAM network can thus achieve rational strategic choice, although it dispenses with the assumption of perfect foresight (i.e., rational expectations). In the special case in which the network stabilizes with the activation of only one strategy in each layer, that pair of strategies is guaranteed to be a Nash equilibrium. Moreover, every pure strategy Nash equilibrium is a stable state in the network. Finally, as starting activation states in the BAM network influence final stable states, cognitively-grounded assumptions regarding the determinants of strategy salience easily resolve issues of equilibrium selection.

The BAM network has previously been used to model reasoning in non-strategic judgment and decision tasks (Bhatia, 2016; see also Howard & Kahana, 2002 for an application of BAM to human memory). In Bhatia's (2016) model, the two layers of the BAM network are assumed to correspond to responses (e.g. hypotheses or choice options) and cues (e.g. evidence or choice attributes). Bidirectional connectivity between these two layers implies that not only do cues activate supported responses, but responses in turn activate the cues that support them. The BAM network is a mathematically tractable variant of the Co3 model analyzed by Spellman, Ullman, and Holyoak (1993) and Holyoak and Simon (1999), and the PCS-DM model analyzed by Glöckner and Betsch (2008) and Glöckner et al. (2014). As with these preexisting models it is able to generate coherence shifts in cue-based judgment and multiattribute choice through the dynamics of the spread of activation in the two layers. Starting point biases formalize the effects of anchors, and activation dynamics generate sequential adjustment in numerical response tasks. Finally starting point biases also capture the effect of reference points on both choice behavior as well as on attribute memory and attention.

The current specification of BAM in game theoretic decision making can be seen as a conceptual extension of Bhatia (2016). In the current paper, the two layers correspond to *self* and *other* strategies. The former are the set of responses that the decision makers choose between, and the latter are the reasons or cues that support these choices. Even though connections can be asymmetric, they nonetheless correspond to support in the same manner as connection weights in Co3, PCS-DM and other constraint satisfaction frameworks.

It is interesting to note that bidirectional reasoning has, in some cases, been argued to be irrational—certainly the effect of beliefs and preferences on the use of evidence and on the evaluation of choice attributes is often considered to be a "bias" in psychological research (Fischhoff & Beyth, 1975; Nisbett & Ross, 1980; Tversky & Kahneman, 1974). In contrast, Thagard (1989) has argued for the value of this type of bidirectional processing in scientific and legal reasoning. In support of Thagard's claims, our work shows that this type of directional process is a core property of rational strategic deliberation. Indeed, strategies for *self* and *other* are only considered to be part of Nash equilibria if they are mutually reinforcing.

Our paper leaves open the question of strategy learning. We have assumed the network is able to encode best responses, but how do decision makers learn these best responses? One solution to this problem may involve a simple form of supervised learning. If, after playing each game (either hypothetically or for real), the decision maker is able to infer both the best response to the strategy that was played by *other*, and *other's* best response to the strategy that was played by *self* ţhen a variant of the perceptron learning rule (that is restricted to binary connections) could over time allow the network to learn the pattern of best response connectivity assumed in our BAM network. This could also be accomplished by reinforcement learning, as the individual considers various possible responses to each strategy. In the long run, reinforcement learning would ensure that the response with the highest payoff would be selected for, and the connection to the best response would be the strongest of all connections. To the extent that the structure of many natural games is learnt, our implementation of strategic reasoning in BAM suggests that optimal strategic choice in games from memory may emerge entirely out of associative processes implemented in a constraint satisfaction network.

Finally, although game theory is most commonly used to describe economic phenomena, such as price setting in markets and bidding in auctions, it also has a large number of applications for the study of cognition and behavior. For example, games such as the traveler's dilemma can help us understand some of the difficulties involved in maintaining cooperation in social settings (Axelrod & Hamilton, 1981). Likewise, rock–paper–scissors can be used to represent evolutionary predator–prey dynamics (Nowak & Sigmund, 2004), and the coordination game provides a perspective for understanding the evolution of language, as different individuals have to agree on the meaning of words (Demichelis & Weibull, 2008). Thus our analysis of the BAM network in this paper has the potential to address core questions regarding human cognition and behavior. We look forward to further integrating psychological research and research on game theoretic decision making in future work.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396.

Basu, K. (1994). The traveler's dilemma: Paradoxes of rationality in game theory. *American Economic Review*, *84*(2), 391–395.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica*, 1007–1028.

Bhatia, S. (2016). The dynamics of bidirectional thought. *Thinking and Reasoning*, *22*(4), 1–46.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Capra, C., Holt, C., Goeree, J., & Gomez, R. (1999). Anomalous behavior in a traveler's dilemma? *American Economic Review*, *89*(3), 678–690.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences*, *26*(2), 139–198.

Crawford, V. P., Gneezy, U., & Rottenstreich, Y. (2008). The power of focal points is limited: even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, *98*(4), 1443–1458.

Demichelis, S., & Weibull, J. W. (2008). Language, meaning, and games: A model of communication, coordination, and evolution. *American Economic Review*, *98*(4), 1292–1311.

Fischhoff, B., & Beyth, R. (1975). I knew it would happen: Remembered probabilities of once—future things. *Organizational Behavior and Human Performance*, *13*(1), 1–16.

Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making*, *3*(3), 215–228.

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition*, *133*, 641666.

Goeree, J. K., & Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *The American Economic Review*, *91*(5), 1402–1422.

Golman, R., & Bhatia, S. (2017). A cognitive model of strategic deliberation and decision making. In *Proceedings of the 39th annual conference of the cognitive science society*.

Golman, R., Bhatia, S., & Kane, P. Dual accumulator model of strategic deliberation and decision making (in preparation).

Harsanyi, J. C., & Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, MA: MIT Press.

Hart, S. (1992). Games in extensive and strategic forms. In *Handbook of game theory with Economic Applications, Vol. 1* (pp. 19–40).

Holyoak, K. J., & Powell, D. (2016). Deontological coherence: A framework for commonsense moral reasoning. *Psychological Bulletin*, *142*(11), 1179–1203.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, *128*(1), 3.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*(3), 295–355.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554–2558.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kosko, B. (1987). Adaptive bidirectional associative memories. *Applied Optics*, *26*, 4947–4960.

Kosko, B. (1988). Bidirectional associative memories. *IEEE Transactions on Systems, Man and Cybernetics*, *18*(1), 49–60.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, *103*, 284–308.

Loomes, G. (2015). Variability, Noise, and Error in Decision Making Under Risk. In *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 658–695).

Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: John Wiley Sons.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: connectionist and dynamical systems approaches to cognition. 14, 348–356.

McClelland, . J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: An account of basic findings. *Psychological Review*, *88*(5), 375.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, *10*(1), 6–38.

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review*, *84*(3), 658–673.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, *85*(5), 1313–1326.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*.

Nowak, M. A., & Sigmund, K. (2004). Evolutionary dynamics of biological games. *Science*, *303*(5659), 793–799.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 1029–1050.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and Gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review*, *1*, 26–53.

Schelling, T. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, *15*(5), 331–336.

Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology*, *86*(6), 814.

Spellman, B. A., Ullman, J. B., & Holyoak, K. J. (1993). A coherence model of cognitive consistency. *Journal of Social Issues*, *4*, 147165.

Stahl, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, *25*(3), 309–327.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, *12*(3) 435–502.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131.