

Modeling Two-Player Games in the Sigma Graphical Cognitive Architecture

David V. Pynadath¹, Paul S. Rosenbloom^{1,2}, Stacy C. Marsella^{1,2}, and Lingshan Li^{1,2}

¹Institute for Creative Technologies, ²Department of Computer Science
University of Southern California, Los Angeles, CA USA

Abstract. Effective social interaction and, in particular, a *Theory of Mind* are critical components of human intelligence, allowing us to form beliefs about other people, generate expectations about their behavior, and use those expectations to inform our own decision-making. This article presents an investigation into methods for realizing Theory of Mind within Sigma, a graphical cognitive architecture. By extending the architecture to capture independent decisions and problem-solving for multiple agents, we implemented Sigma models of several canonical examples from game theory. We show that the resulting Sigma agents can capture the same behaviors prescribed by equilibrium solutions.

1 Introduction

Social interaction is a critical function of human intelligence, allowing us to make effective decisions when interacting with other people. We form beliefs about others, use those beliefs to generate expectations about their behavior, and use those expectations to inform our own behavior. In doing so, we attribute an intelligence like our own to other people as well. This cognitive capacity for *Theory of Mind* distinguishes social interaction from the decision-making we do in isolation [18]. We therefore expect that a system capable of artificial general intelligence (AGI) would provide natural support for Theory of Mind. Previous computational approaches have successfully captured aspects of this cognitive capacity, including modeling the beliefs of others [4, 6], predicting behavior [1], and factoring behavior predictions into decision-making [3, 13].

We are interested here in how all of these Theory of Mind capabilities may be realized within *Sigma* (Σ), a nascent *cognitive system*—an integrated computational model of intelligent behavior—that is grounded in a *cognitive architecture*, a model of the fixed structure underlying a cognitive system [11]. In prior work, we have demonstrated this architecture’s ability to support multiple cognitive capabilities, such as problem solving [14], mental imagery [16], and learning [15]. Here, our goal is to extend this existing integration to the critical cognitive capability of Theory of Mind, while shedding additional light on both Sigma and the architectural basis of Theory of Mind.

To that end, we explore the degree to which existing capabilities within Sigma’s cognitive architecture already provide the necessary support for Theory of Mind. We draw on canonical examples from the game theory literature as test cases for driving the need for reasoning about other agents [2]. These games provide a natural setting for Theory of Mind, because agents seeking to perform well must take each other’s action

selection into account. Across these games, we examine the conditions under which the Sigma-generated outcomes conform to those prescribed by game-theoretic analyses.

We also identify the need for new capabilities within Sigma to address these games and present the extensions we implemented to provide them. Sigma turns out to support two distinct approaches to Theory of Mind, based respectively on automatic versus controlled processing [17] or, alternatively, System 1 versus System 2 thinking [7]. Both modes of reasoning are possible without any change to the underlying architecture, providing a novel unification of the two forms of Theory of Mind within a single cognitive system. By successfully realizing Theory of Mind reasoning and combining it with Sigma's other cognitive capabilities (e.g., learning), we now open up novel lines of investigation into more general mechanisms for social reasoning.

2 Sigma

Sigma's cognitive architecture is built upon *graphical models* [8]. Graphical models provide a general computational technique for efficient computation with complex multivariate functions by leveraging forms of independence to: decompose them into products of simpler functions; map these products onto graphs; and solve the graphs via, for example, message passing or sampling methods. They are particularly attractive as a basis for broadly functional, yet simple and theoretically elegant cognitive architectures, because they provide a single general representation and inference algorithm for processing symbols, probabilities and signals. The cognitive architecture leverages this generality through a core knowledge structure—the *conditional*—that provides a deep blending of the forms of conditionality found in both rules and probabilistic networks.

Sigma's long-term memory comprises a set of conditionals, which are jointly compiled into a single *factor graph* [9] at the level below. Memory access occurs by passing messages in this graph, via the *summary product algorithm* [9], until quiescence; that is, until there are no more messages to send. Each message is an n -dimensional piecewise linear function that is defined over an array of rectilinear regions. These piecewise linear functions can approximate arbitrary continuous functions as closely as desired, as well as be restricted to represent both discrete probability distributions and relational symbol structures. Working memory consists of a set of peripheral nodes in the graph that provide fixed *evidence* during solution of the long-term-memory graph.

This way of viewing the Sigma cognitive system divides it into three layers: one for the graphical models, one for the (Sigma) cognitive architecture, and one for the knowledge and skills encoded via conditionals on top of the cognitive architecture. However, the Sigma cognitive architecture also embodies its own hierarchy of three layers that is central to realizing Theory of Mind. Modeled after the Soar architecture [10], the Sigma architecture comprises a *reactive layer* that acts as the inner loop for a *deliberative layer* that acts as the inner loop for a *reflective layer*. The reactive layer simply consists of quiescent memory access. In Soar, this form of *knowledge search* provides a parallel associative retrieval of information stored in rules. In Sigma it searches over the more general representation provided by conditionals and factor graphs.

Sigma's deliberative layer provides a model of sequential action [14], during which operators are selected and applied in service of achieving goals (or of maximizing util-

ities). In Sigma, as in Soar, operator selection and application are both mediated by the reactive layer, with the core *cognitive cycle* becoming quiescent memory access followed by a decision. In general, the distinction between reactive and deliberative layers maps onto two well-known distinctions in human cognition: automatic versus controlled [17] and System 1 versus System 2 [7].

The reflective layer determines how metalevel processing—including *problem space search*—occurs when resolving *impasses* in deliberative processing. In Sigma, three kinds of impasses can occur when attempting to make a decision. A *none* impasse occurs when there are no candidate operators for selection. A *tie* impasse occurs when there are multiple candidate operators, but insufficient knowledge to choose among them. A *no-change* impasse occurs when an operator is selected but no state change results. When processing begins, there is only a base-level state (state 0) active. When an impasse happens in a state, a structure representing it is added to working memory and a new state with a number one higher is added. Whenever an impasse occurs for a state, all existing higher numbered states are deleted, with just the one new state then added. Thus, if impasses occur simultaneously for multiple states, only the lowest numbered impasse goes into effect, as all higher numbered ones are for states that are deleted.

An impasse, along with its resulting state—and all higher numbered states (which ultimately only exist in service of this lower numbered impasse)—will also go away when it is no longer valid. This can happen because the impasse is resolved: the appearance of a candidate operator resolves a none impasse, the selection of an operator resolves a tie impasse, and a change in the state resolves a no-change impasse. It can also happen because the impasse changes, such as when the appearance of multiple undifferentiated operators shifts the impasse from none to tie. Communication across states in the metalevel hierarchy occurs via the *affine transforms* that were earlier implemented in Sigma in service of modeling mental imagery [16]. Translation along the state dimension moves information to higher numbered states (to initialize metalevel states) and to lower numbered states (to return results from reflective processing).

3 Single-Stage Simultaneous-Move Games

We begin our investigation by modeling single-stage simultaneous-move games within Sigma. Two agents, *A* and *B*, choose from two operators: *C* and *D* (typically, *cooperate* and *defect*, respectively). After the players simultaneously reveal their operator selection, the corre-

sponding entry in the *payoff matrix* yields the outcome for *A* and *B*. For example, if *A* cooperates and *B* defects, then Table 1 specifies that *A* gets r_{ACD} and *B* gets r_{BCD} .

The typical analysis of such games revolves around *Nash equilibria*. In such an equilibrium, each agent is unable to increase its reward by *unilaterally* changing its chosen strategy. For example, if $r_{BCC} > r_{BCD}$ and $r_{BDC} > r_{BDD}$, then agent *B* will choose to cooperate, as it yields a higher payoff, regardless of what agent *A* does. Knowing that *B* will cooperate, agent *A* will also choose to cooperate if $r_{ACC} > r_{ADC}$. If both

<i>B</i> →		<i>C</i>	<i>D</i>
<i>A</i>	<i>C</i>	r_{ACC}, r_{BCC}	r_{ACD}, r_{BCD}
	<i>D</i>	r_{ADC}, r_{BDC}	r_{ADD}, r_{BDD}

Fig. 1: Payoff matrix for two-player, simultaneous-move game

conditions hold, the situation where both agents cooperate constitutes a Nash equilibrium, because neither has incentive to defect. In this investigation, we consider only such *pure-strategy* equilibria, where the agents’ choices have no stochastic component.

There are at least three different ways the independent decision-making processes of A and B could conceivably be represented within Sigma. Multiple agents could be represented in Sigma as multiple instances of the whole system, or as multiple graphs (and associated decision making) within a single instance of Sigma, or by adding an agent dimension to functions within a single graph. For this work, we chose the third approach, so as to maximize the possibility of sharing functional structure among multiple agents, and thus to minimize the cost of doing complex Theory of Mind reasoning. In addition to adding an agent dimension to functions, this involved simply extending the decision procedure to make independent decisions for each agent.

To model such games in Sigma, we require four conditionals to completely capture the generation of the two agents’ expectations and decisions. Two conditionals ($AB \rightarrow AA$ and $BA \rightarrow BB$ in Figure 2) specify the value of each agent’s operator selection as a function of its expectations about the other agent’s selection. For example, the function at PAB would specify that if B chooses

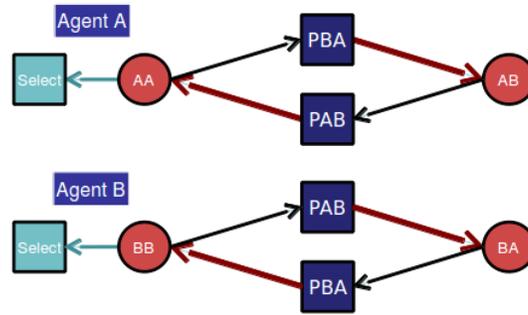


Fig. 2: Factor graph generated for one-shot games

$o_B \in \{C, D\}$, then the value to A of choosing $o_A \in \{C, D\}$ would be $r_{A o_A o_B}$. Agent B has an analogous conditional ($BA \rightarrow BB$) regarding the impact of its expectation of A ’s choice. The remaining two conditionals ($AA \rightarrow AB$ and $BB \rightarrow BA$) capture A ’s and B ’s expectations about each other’s action values in response to their own choices.

When executing this model, Sigma begins by computing values over each agent’s options (AA and BB in Figure 2), first assuming a uniform distribution over the other agent’s possible moves (AB and BA). The resulting values are normalized into a probability distribution over the deciding agent’s selection, which then feeds, in turn, into the other agent’s conditional ($AA \rightarrow AB$ and $BB \rightarrow BA$). The values go around this loop several times before converging. By using a linear transformation of action values into probabilities, each agent assigns a nonzero probability that the other does *not* choose a best response, possibly deviating from a Nash equilibrium. However, if agents select operators stochastically according to this distribution (instead of deterministic maximization), then the agents would be playing a *quantal response equilibrium*, as their actions would be a best response to the probability of each other’s choices [12].

Perhaps the most famous example of a single-stage, simultaneous-move game is the *Prisoner’s Dilemma* (Table 1a). From the payoff matrix of Table 1a, we can see that each agent is better off by choosing to *defect*, regardless of what the other agent does. Thus, there is a single pure-strategy Nash equilibrium, where both agents choose D , even though both agents would be better off if they both chose C .

		$B \rightarrow$	
		C	D
A	C	0.3, 0.3	0.1, 0.4
	D	0.4, 0.1	0.2, 0.2

(a) Prisoner's Dilemma

		$B \rightarrow$	
		C	D
A	C	0.3, 0.3	0.0, 0.1
	D	0.1, 0.0	0.1, 0.1

(b) Stag Hunt

		$B \rightarrow$	
		C	D
A	C	0.1, 0.1/0.4	0.2, 0.4/0.1
	D	0.3, 0.1/0.4	0.1, 0.4/0.1

(c) Dominant B strategy

		$B \rightarrow$	
		C	D
A	C	0.1, 0.2	0.2, 0.1
	D	0.2, 0.1	0.1, 0.2

(d) Matching pennies

Table 1: Payoff matrices for one-shot games

When we use Table 1a's payoff matrix, Sigma requires 401 messages to compute the agents' action selection values. The resulting values favor defecting over cooperating, 0.572 vs. 0.428, for both agents, matching the Nash equilibrium. We are thus able to use a single graph to model both agents' individual, self-interested perspectives, as opposed to finding the C - C outcome that is socially optimal from a global perspective.

The *Stag Hunt* (Table 1b) is an example of a coordination game, because any outcome where the two agents choose the same action is a Nash equilibrium. Intuitively, if both agents cooperate to hunt a stag, they will succeed and receive the maximum payoff. However, any player who defects to hunt a hare will succeed, albeit for a smaller payoff. An agent who attempts to hunt the stag unassisted (i.e., chooses to cooperate when the other agent defects) will fail and receive no payoff. Thus, while this game shares the same D - D equilibrium as the Prisoner's Dilemma, it also supports a C - C equilibrium, which is also the socially optimal outcome.

In the Stag Hunt, Sigma requires 681 messages to compute a distribution where both agents favor cooperating over defecting with values of 0.573 vs. 0.427. Our Sigma model thus finds the socially optimal Nash equilibrium. The values are similar to those for the Prisoner's Dilemma, although the cycle requires significantly more messages.

In the Prisoner's Dilemma, defecting is dominant for each agent regardless of the other agent's choice. In the Stag Hunt, agents greedily seeking the highest potential payoff might also find the C - C equilibrium accidentally. We implemented the game of Table 1c to make sure that our Sigma agents are properly considering their opponents' decisions. There are two variations for B 's payoffs, separated by a "/". With the left payoffs, D is a dominant action for B , while with the right, C is dominant. In the former case, C is A 's best response, while in the latter, D is the best response. Thus, we can generate two different optimal decisions for A by changing only B 's payoffs.

For both versions of this game, Sigma computes a distribution over actions after a total of 226 messages. B 's values are symmetric across both versions, with the dominant action being preferred 0.708 vs. 0.292. When B prefers C , A prefers D 0.595 vs. 0.405, while when B prefers D , A prefers C 0.511 vs. 0.489. Thus, agent A makes the equilibrium-specified decision under both cases, while the higher payoff for D in Table 1c shows up in the stronger bias in the former case.

Unlike the previous games, the *Matching Pennies* game (Table 1d) has no pure-strategy Nash equilibrium. If both agents choose the same action, B gets a higher payoff; otherwise, A gets the higher payoff. To see that no pure-strategy equilibrium exists, consider the case if A chooses C , so that B chooses C , so that A chooses D , so that B chooses D , so that A chooses C again, returning us to the beginning of this sequence.

In this game, Sigma finishes computation after only 151 messages with each agent having a value of 0.5 for both operators. With this distribution over values, neither agent has a preference over its operator selection. Given the lack of any equilibrium in this game, it makes sense for the agents to be indifferent among their options.

Sigma is thus able to find the equilibrium strategies through graphical message-passing, without requiring the logical, case-by-case reasoning that is required in general. Of course, as already described, the graph is not guaranteed to find the equilibrium strategy in *all* games. There is a potential relationship between the quantitative values computed by Sigma and the frequency with which people play equilibrium strategies in the same games, but studying that relationship is beyond the scope of this investigation.

4 Sequential Games: The Ultimatum Game

In Section 3's games, both players reveal their moves simultaneously. In contrast, the agents in *sequential* games take turns choosing their actions. In this section, we examine the *ultimatum game* [5], where agent A starts with a fixed amount of money (3 in our example), and offers a portion of it to B , who then accepts or rejects the offer. If B accepts, it receives the offered amount, while A keeps the remainder. However, if B rejects, both agents get 0 (Figure 3 shows the game tree).

The game's sequential nature leads to a different strategy structure than in simultaneous-move games. While agent A 's strategy is still a single choice (the amount to offer), agent B 's strategy is now a function from A 's possible offers to a binary acceptance/rejection. There are many Nash equilibria within this strategy space, including one if A offers 0 and B accepts any offer. B has no incentive to reject an offer of 0, and if B accepts any offer, then A has no incentive to offer more. At the other extreme, there is an equilibrium where A offers everything to agent B and B rejects if offered anything less than everything. Again, neither agent has reason to unilaterally deviate.

This highly divergent space of Nash equilibria leads us instead to the stricter concept of *subgame perfection*, where the overall equilibrium strategy is also an equilibrium strategy over any subgame in the game tree. For example, the second Nash equilibrium mentioned is *not* a subgame perfect equilibrium, because it is not an equilibrium strategy for B to (for example) reject in the subtree rooted after A offers only 1. Therefore, it is not *credible* for B to reject any offer less than everything. We can determine the subgame perfect equilibrium strategy by backward induction from the end payoffs. If A offers any amount more than 0, B 's best response is to accept the offer, because any positive offer is better than the 0 it would get by rejecting. Out of these positive offers,

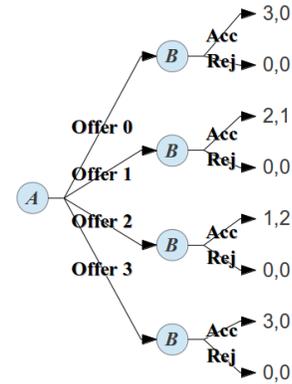


Fig. 3: Extensive form of ultimatum game

A prefers offering the minimal amount of 1, which forms a subgame perfect equilibrium strategy in combination with B 's acceptance of any nonzero offer.

We present two contrasting methods of decision-theoretic Theory of Mind processing for a Sigma agent to use in such sequential games. In the first (Section 4.1), the agent's reasoning takes place within a single cognitive cycle, as in Section 3's approach to single-stage games. In other words, the agent's decision is the outcome of memory access (i.e., message passing) within Sigma's reactive layer. In contrast, our second method (Section 4.2) leverages Sigma's deliberative and reflective layers to perform combinatoric search across multiple problem spaces, as the agents reach impasses (e.g., when there are no alternatives available for operator selection). The ability of humans to solve problems in either fashion is well documented, but this flexibility is rare in AGI systems, and previously non-existent for Theory of Mind.

4.1 Automatic Processing in the Ultimatum Game

We label the first method as *automatic* processing, as it is represented directly at Sigma's reactive layer, via a *trellis* graph that passes utilities at terminal states backwards over transition functions that represent the agents' actions. Figure 4 shows this graph for the ultimatum game, where *offer* represents A 's possible offers, and *accept* represents B 's binary decision of acceptance as a function of A 's offer. We

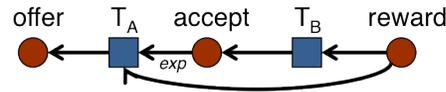


Fig. 4: Trellis-shaped factor graph for automatic processing in the ultimatum game

translate the money the agents receive at the end into the *reward* function according to how they value money. In the example implemented here, both agents value 0 at .1, 1 at .4, 2 at .7 and 3 at 1. If we are considering A 's model of B , then (A 's model of) B 's evaluation of the outcomes is converted into (A 's model of) B 's utilities for its accept and reject operators, given A 's offers, by passing these utilities backwards over the transition function for (A 's model of) B 's actions (T_B). In other words, there is a Sigma conditional that passes the values over B 's ending money (m_B) to B 's *accept* operator, where accepting an offer of q implies that $m_B = q$, while rejecting any offer implies that $m_b = 0$.

While the values computed here are sufficient for B 's operator selection through utility maximization, we must then convert these utilities into a probability distribution for A 's model of B 's choices. If A uses these values as probabilities directly, then the resulting distribution will be linear in those utilities (as in Section 3). Reinforcement learning more commonly uses a *softmax* function to convert values over operators, O , into probabilities, $\text{Pr}(O)$, according to $\text{Pr}(O = o) \propto \exp(V(o)/\tau)$, for some temperature parameter, τ , and after normalization. Game theory's *logit* form of a *quantal response equilibrium* (LQRE) uses a similar form to map values into probabilities [12].

We can incorporate the same concept by exploiting Sigma's support of *Boltzmann selection*, allowing an agent to choose operators probabilistically according to an exponential transformation of utilities, as in the softmax and LQRE formulations. This capability had been used previously to encourage exploration when Sigma agents are performing reinforcement learning. The multiagent context is slightly different, in that we need to transform (A 's model of) B 's utilities in a corresponding manner to yield

B 's choice probabilities. To enable this, the cognitive language was extended to allow messages coming from particular working memory functions to be scaled exponentially (and normalized), using the architecture's support for Boltzmann distributions.

Evaluations for A 's offer operators are then computed by passing the products of A 's own utilities for B 's choices and its model of B 's probabilities for these choices backwards over the transition function for A 's own actions (T_A), while summarizing out—via integration across—(A 's model of) B 's choices. Both the product and summarization required here are computed naturally by the summary product algorithm underlying this processing.

All together, the processing involves five conditionals that define: (1) the utility functions of the agents; (2) the transition function for A 's offer operators; (3) (A 's model of) the transition function for B 's accept operators; (4) the exponential transformation from (A 's model of) B 's operator utilities to its operator selection probabilities; and (5) the product of (A 's model of) B 's operator probabilities and A 's utilities. A total of 94 messages are required to compute a distribution over the agents' operators. B 's contingent strategy has normalized values of $\langle 0.500, 0.800, 0.875, 0.909 \rangle$ for accepting the corresponding offer of 0, 1, 2, or 3 from A . After passing through the Boltzmann distribution and A 's transition function, A arrives at values of $\langle 0.315, 0.399, 0.229, 0.057 \rangle$ for offering 0, 1, 2 or 3. We thus arrive at the subgame perfect equilibrium of A offering 1 and B accepting nonzero offers. Although Sigma's reactive layer is modeled after Soar's, the extension to graphical models, and the resulting ability to perform probabilistic reasoning over trellises, is central to decision-theoretic Theory of Mind here.

4.2 Controlled Reasoning in the Ultimatum Game

Sigma's ability to engage in problem space search across metalevel states can be leveraged to provide our second method for Theory of Mind: a *controlled* form of decision-theoretic reasoning. This method of Theory of Mind reasoning required an extension of Sigma's impasse detection and processing to multiple agents. Simple single-agent approaches that used mechanisms related to affine transforms to modify the state dimension had to be replaced with more elaborate mechanisms that operated independently on segments of functions corresponding to individual agents.

Figure 5 illustrates this multiagent mechanism for the ultimatum game *without* the trellis in long-term memory. Four undifferentiated offer operators—for 0 through 3—are acceptable, and a tie impasse occurs (at the top-left node of Figure 5). At the next higher state, evaluation (meta-)operators are acceptable, one for each offer operator, and made *best*—that is, receive a rating of 1—so that one is chosen at random. A no-change results on this (meta-)operator since direct knowledge about how to evaluate its operator is unavailable. At the next higher state, the operator to be evaluated is made best, and thus selected and applied, yielding a specific offer. A second-level tie impasse then occurs on (A 's model of) B 's choice of accept versus reject, yielding a second round of choices among evaluation (meta-)operators. At the next higher state, accept or reject is selected according to the evaluation (meta-)operator selected. It is then applied, yielding the resulting monetary outcomes for both A and B , plus the utilities of these outcomes (at the bottom of the leftmost tree of Figure 5).

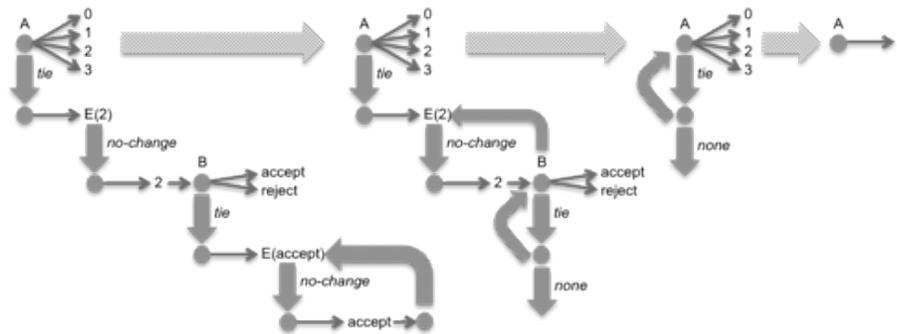


Fig. 5: Progression of problem spaces during controlled reasoning in ultimatum game

Returning results up a level enables terminating the evaluation (meta-)operator and selecting the other one. Once both are terminated, a none impasse occurs, as there are no more evaluation (meta-)operators (the middle tree of Figure 5). This is the signal to combine the results across B 's choices for the offer operator being evaluated for A , terminating the higher evaluation (meta-)operator, and enabling other offer operators to be evaluated in the same manner. Once a none impasse occurs here, these evaluations are returned, enabling A to select the best offer (as in the rightmost tree of Figure 5).

The distribution over the four offers calculated in this fashion is $\langle .0550, 0.699, 0.400, 0.100 \rangle$, which when normalized to $\langle 0.314, 0.400, 0.229, 0.057 \rangle$ again yields the subgame perfect equilibrium, with a distribution that is comparable to that computed reactively via the trellis. The knowledge here comprises 19 conditionals, 3 of which are general conditionals for evaluation (meta-)operators—corresponding to what are called *default rules* in Soar—and 16 specific to this task. The processing occurs across 72 cognitive cycles, each involving on average 868 messages.

5 Conclusion

We have shown how Sigma can capture different types of Theory of Mind in several multiagent domains. The success of this endeavor relied on extensions to the architecture to be able to maintain the individual reasoning of multiple agents, whether within a single level at the reactive layer, or across nested problem spaces. The reactive approach in Sigma provides a more efficient compiled form of the calculation, gaining a factor of 6334 in speedup in the ultimatum game, where automatic processing takes 0.02s on a laptop, while the controlled version takes 126.69s. In Soar, chunking is a learning mechanism that compiles reflective problem solving into more efficient reactive rules. Although Sigma does not yet embody a comparable learning mechanism, we do have initial ideas for a generalization of chunking that could potentially compile the sequential problem solving of Section 4.2 into the reactive trellises of Section 4.1.

In addition to using Theory of Mind to motivate necessary extensions to Sigma, we also open up the possibility of exploiting Sigma's other cognitive capabilities in the service of Theory of Mind. For example, all of the models presented here gave the agents complete information, so that the payoffs of both agents were common knowledge. It

is straightforward to relax that assumption and give the agents uncertainty regarding each other's payoffs. If we model that uncertainty within Sigma, we can potentially apply its learning capability [15] to allow the agents to learn models of each other in a repeated-game setting. We thus hope that continued progress in this direction can lead to a synergy that can generate insights into both the requirements of AGI in multiagent contexts and general-purpose mechanisms for Theory of Mind.

Acknowledgements This work has been sponsored by the Office of Naval Research and the U.S. Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government.

References

1. Baker, C.L., Saxe, R., Tenenbaum, J.B.: Action understanding as inverse planning. *Cognition* **113**(3) (2009) 329–349
2. Fudenberg, D., Tirole, J.: *Game Theory*. MIT Press (1991)
3. Goodie, A.S., Doshi, P., Young, D.L.: Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making* **25**(1) (2012) 95–108
4. Goodman, N.D., Baker, C.L., Bonawitz, E.B., Mansinghka, V.K., Gopnik, A., Wellman, H., Schulz, L., Tenenbaum, J.B.: Intuitive theories of mind: A rational approach to false belief. In: *Proceedings of the Conference of the Cognitive Science Society*. (2006) 1382–1387
5. Güth, W., Schmittberger, R., Schwarze, B.: An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization* **3**(4) (1982) 367–388
6. Hiatt, L.M., Trafton, J.G.: A cognitive model of theory of mind. In: *Proceedings of the International Conference on Cognitive Modeling*. (2010)
7. Kahneman, D.: *Thinking, fast and slow*. Farrar, Straus and Giroux (2011)
8. Koller, D., Friedman, N.: *Probabilistic graphical models: principles and techniques*. MIT Press (2009)
9. Kschischang, F.R., Frey, B.J., Loeliger, H.A.: Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory* **47**(2) (2001) 498–519
10. Laird, J.E.: *The Soar Cognitive Architecture*. MIT Press (2012)
11. Langley, P., Laird, J.E., Rogers, S.: Cognitive architectures: Research issues and challenges. *Cognitive Systems Research* **10**(2) (2009) 141–160
12. McKelvey, R.D., Palfrey, T.R.: Quantal response equilibria for normal form games. *Games and Economic Behavior* **10**(1) (1995) 6–38
13. Pynadath, D.V., Marsella, S.C.: PsychSim: Modeling theory of mind with decision-theoretic agents. In: *Proceedings of the International Joint Conference on Artificial Intelligence*. (2005) 1181–1186
14. Rosenbloom, P.S.: From memory to problem solving: Mechanism reuse in a graphical cognitive architecture. In: *Proceedings of the Conference on Artificial General Intelligence*, Springer (2011) 143–152
15. Rosenbloom, P.S.: Deconstructing reinforcement learning in sigma. In: *Proceedings of the Conference on Artificial General Intelligence*. (2012)
16. Rosenbloom, P.S.: Extending mental imagery in sigma. In: *Proceedings of the Conference on Artificial General Intelligence*. (2012)
17. Schneider, W., Shiffrin, R.M.: Controlled and automatic human information processing: 1. detection, search, and attention. *Psychological Review* **84** (1977) 1–66
18. Whiten, A., ed.: *Natural Theories of Mind*. Basil Blackwell, Oxford, UK (1991)