Estimating the use of higher-order theory of mind using computational agents

Harmen de Weerd, Denny Diepgrond, Rineke Verbrugge
Institute of Artificial Intelligence, University of Groningen

Abstract

When people make decisions in a social context, they often make use of *theory of mind*, by reasoning about unobservable mental content of others. For example, a pedestrian who wants to cross a street behaves differently depending on whether or not he believes that the driver of an oncoming car has seen him or not. People can also reason about the theory of mind abilities of others, leading to recursive thinking of the sort 'I think that you think that I think...'. Previous research suggests that this ability may be especially effective in simple competitive settings. In this paper, we use a combination of computational agents and Bayesian model selection to determine to what extent people make use of higher-order theory of mind reasoning in a particular competitive game known as matching pennies. We find that while many children and adults appear to make use of theory of mind, participants are also often classified as using a simpler strategy based only on the actions of the directly preceding round. This may indicate that human reasoners do not primarily use their theory of mind abilities to compete with others.

1 Introduction

In social interactions, people often reason about the beliefs, goals, and intentions of others. People use this so-called *theory of mind* [14] or *mentalizing* to understand why others behave the way they do, as well as to predict the future behavior of others. People can even use their theory of mind to reason about the way others make use of theory of mind. For example, people make use of *second-order theory of mind* to understand a sentence such as "Alice *knows* that Bob *knows* that Carol is throwing him a birthday party", by reasoning about what Alice knows about what Bob knows.

The human ability to make use of higher-order theory of mind is especially apparent in story comprehension tasks. Adults perform much better than chance on story comprehension questions that require up to fourth-order theory of mind [10, 18]. Interestingly, experimental evidence shows that people have more difficulty applying their theory of mind abilities in strategic games. In these settings, individuals are typically found to reason at low orders of theory of mind and are slow to adjust their level of theory of mind reasoning to more sophisticated opponents [1, 7, 8, 23]. However, some empirical research suggests that the use of theory of mind by participants can be facilitated by context [2, 11], setting [7, 20, 21], and training [12].

Both empirical studies and simulation studies suggest that the ability to make use of higher-order theory of mind may be particularly useful in simple competitive games [2–5, 7, 22]. However, in these simple competitive settings, it is difficult to distinguish participants who make use of theory of mind from participants who rely on simpler, behavior-based strategies. In this paper, we use a combination of computational agents and Bayesian model selection, introduced by [17], to estimate strategy use of participants in simple strategic games. Our goal is to test the effectiveness of this estimation procedure, as well as to determine to what extent human participants make use of theory of mind in simple competitive games. To do so, we apply this method on two empirical studies in which participants play a simple game known as matching pennies.

In the first study, Devaine, Hollard, and Daunizeau [2] let participants play against Bayesian theory of mind opponents. Participants played a hide-and-seek task, in which they searched for an opponent who was hidden in one of two possible hiding locations, and a casino task, in which participants had to choose one of

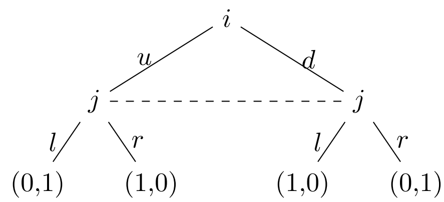

Figure 1: The hide-and-seek game [2] and the sender-receiver game [16] share a common underlying structure. First, player i chooses to perform either action u or action d. Without knowing what choice player i has made, player j then chooses to perform either action l or r. This results in an outcome (π^i, π^j), where player i obtains payoff π^i and player j obtains outcome π^j.

two slot machines to play against. Importantly, these two tasks differed only in the cover story presented to the participants. Both games had the exact same structure and the exact same computer opponent.

In the second study, Sher, Koenig, and Rustichini [16] let children between the ages of 3 and 9 play a sender-receiver game. This game involved two boxes, one of which hid a piece of candy, while the other box hid a rock. Only the sender was told which box contained the candy. The sender was then asked to point at one of the boxes, after which the receiver could choose one of the two locations. If the receiver selected the box with the candy, the receiver could keep the candy. Otherwise, the sender would get the candy.

In both these studies, human participants played against an opponent that followed a known and fixed strategy. In Devaine et al. [2], participants played against software agents that followed a theory of mind strategy, while Sher et al. [16] let children play against a confederate who always selected the action that would have won in the last round. In this paper, we estimate the level of theory of mind reasoning of both the participants as well as their opponents. This allows us to both validate the estimation method, by comparing our estimation results to the known strategies, and estimate the extent of human theory of mind use in simple competitive games.

The remainder of this paper is structured as follows. In Section 2, we present the details of the game that participants play in the studies of Devaine et al. and Sher et al. Section 3 outlines the estimation method and agent strategies that we consider in our estimation. The results of the estimation are presented in Section 4. In Section 5, we discuss these results and suggest directions for future research.

2 Matching pennies

Matching pennies is a simple two-player game in which two players, i and j, independently select one of two possible actions. Figure 1 shows an extensive form representation of the game. In this figure, the actions resulting in the outcomes (u, l) and (d, r) are considered 'matched', which results in player j winning. In the remaining outcomes (u, r) and (d, l), the actions are 'mismatched', and player i wins.

The hide-and-seek game [2] and the sender-receiver game [16] share the underlying structure of matching pennies. Both games start with player i selecting to perform either action u or action d. Afterwards, without knowing the choice of player i, player j decides whether to perform action l or action r. In the hide-and-seek game [2], player i is the hider that hides at location 0 (action u) or at location 1 (action d). Player j is the seeker that can either search location 0 (action l) or location 1 (action r). In the sender-receiver game [16], player i is the sender who points at the hidden candy (action u) or at the hidden rock (action d). Player j is the receiver who either selects the box pointed out by the sender (action l) or the other box (action r).

The unique Nash equilibrium in matching pennies is for both players to randomly select one of their actions to play [see, for example, 9]. That is, unless both players play randomly, at least one of the players has an incentive to change their behavior. Human participants are known to have difficulties generating random sequences [15, 19]. Participants playing repeated matching pennies games therefore typically deviate from playing the Nash equilibrium. This may encourage participants to try and take advantage of their opponent's deviations. In Section 3, we describe a way to determine what strategies participants use.

3 Estimation method

To determine to what extent human participants make use of theory of mind when playing matching pennies, we make use of a technique known as group-level random-effects Bayesian model comparison (RFX-BMS), introduced by Stephan et al. [17]. Random-effects Bayesian model selection treats strategies as random effects that can vary among participants, and which occur with fixed and unknown population frequencies. That is, unlike fixed-effects Bayesian model selection, we do not assume that there is one strategy that best describes the actions of all participants. Instead, we define a number of strategies that participants may use, including both theory of mind strategies and heuristic strategies. Each of these strategies s generates pieces of evidence $p(y|s)$ representing the probability that following strategy s results in some observed data y. This allows random-effects Bayesian model selection to estimate the relative frequencies of these strategies in the general population of participants.

To test whether participants make use of theory of mind while playing matching pennies, we compare the observed behavior of participants with the predicted behavior of computational agents following different strategies. In addition to these theory of mind strategies, we also consider a number of heuristic strategies. These heuristic strategies do not have an observable internal state. Instead, these strategies only react to the most recently observed behavior. This technique has previously been used by Devaine et al. [2] to estimate the effect of framing on theory of mind use in participants. Instead, we compare theory of mind use across different opponent strategies. Unlike Devaine et al., we also attempt to identify known strategies through Bayesian RFX-BMS estimation. This allows us to determine to what extent Bayesian RFX-BMS estimation can accurately distinguish theory of mind strategies from heuristic strategies.

In the following subsections, we discuss each of the strategies included in the analysis in detail. To avoid confusion, for the remainder of this paper we refer to player i as if she were female, while we refer to player j as if he were male.

3.1 Heuristic strategies

To determine whether participant behavior is better explained by simple heuristics or by the use of theory of mind, we consider three heuristic strategies. The heuristics we describe here do not make use of an internal state. Instead, these strategies respond to actions observed in the previous round of play only. Each strategy is parameterized with a single parameter $\lambda \in [0, 1]$, which we will refer to as the learning speed. Note that although we use the same symbol for each strategy, λ has a different interpretation in each strategy.

The first heuristic strategy we consider is the *biased* strategy. This strategy does not react to the behavior of either player, but instead selects each action with a fixed probability. Specifically, a biased player i selects action u with some individual bias probability λ and action d with probability $(1-\lambda)$. Similarly, when player j follows a biased strategy, he selects action l with probability λ and action r with probability $(1-\lambda)$.

We also consider an *other-regarding* strategy. A player following this strategy selects with probability λ the action that would have resulted in the player winning the previous round. As a result, an other-regarding agent selects the action that would have lost the previous round with probability $(1-\lambda)$. For example, suppose that in the previous round of the game, player i has chosen action u and player j has selected action l. Since this means that player j has won the previous round, an other-regarding player i would select action d with probability λ, while an other-regarding player j would repeat his previous choice l with probability λ.

Finally, we consider a *self-regarding* strategy. A self-regarding player repeats the action performed in the previous round with probability λ and selects the opposite action with probability $(1-\lambda)$. For example, if a self-regarding player i chooses action u, her next action has probability λ of being u and probability $(1-\lambda)$ of being d. This strategy is meant to model players that attempt to play randomly, but either switch between actions too often ($\lambda < 0.5$) or switch too little ($\lambda > 0.5$).

Note that each of these three heuristic strategies is equivalent to the Nash equilibrium strategy when the parameter λ equals 0.5. In this case, each strategy randomly chooses one of the two options to play. In addition to these three heuristic strategies, we also include the Nash equilibrium strategy as a separate parameter-less strategy in our analysis of player strategies.

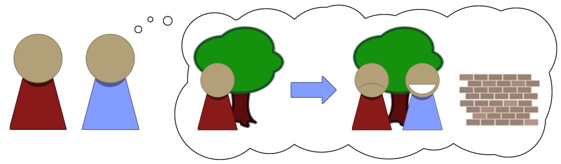

Figure 2: If the blue agent j is a ToM_0 seeker, his beliefs concerning the hiding location of the hider i are based on her previous behavior. If the red agent i has typically hidden herself behind the tree, the blue agent j believes it is most likely that she will hide behind the tree again, and choose to search for her there.

3.2 Zero-order theory of mind

In addition to the heuristic strategies described in the previous subsection, we also consider the possibility that participants apply a strategy based on the use of theory of mind. Note that the theory of mind strategies we describe here implement our theory of mind (*ToM*) agents that we studied in the game of rock-paper-scissors [22]. These agents differ from the Bayesian theory of mind agents used in the study of Devaine et al. [2]. This means that our random-effects Bayesian model selection does not include the exact strategies used by the computational theory of mind agents. By using a different agent model to estimate the use of theory of mind than was used to generate agent behavior in the hide-and-seek experiment, we demonstrate that Bayesian RFX-BMS estimation [17] can estimate the level of theory of mind reasoning of a player, even if that player's strategy deviates from our specification.

A zero-order theory of mind (ToM_0) agent is a goal-directed agent that is unable to represent or reason about mental content. Instead, the ToM_0 agent makes predictions about the behavior of the opponent based only on previously observed behavior. Figure 2 shows an example of this process in the hide-and-seek game. In this example, previous behavior of the red hider agent (agent i) leads the blue ToM_0 seeker agent (agent j) to believe that she has hidden herself behind the tree. As a result, the ToM_0 seeker j believes that he should look for his opponent at the same location.

A ToM_0 agent forms zero-order beliefs $b^{(0)}$ so that $b^{(0)}(a)$ represents the probability that the agent assigns to the event of its opponent playing some given action a. For example, if player i is a ToM_0 agent, she believes that opponent j will play action l with probability $b^{(0)}(l) = 1 - b^{(0)}(r)$. Using these beliefs, the ToM_0 agent can calculate the expected value of playing a given action. For example, the expected value $EV^{(0)}(u)$ that ToM_0 agent i assigns to playing action u is

$$EV_i^{(0)}(u; b^{(0)}) = \sum_{a^j \in \{l,r\}} b^{(0)}(a^j) \cdot \pi^i(u, a^j) = b^{(0)}(r) = 1 - b^{(0)}(l). \tag{1}$$

The ToM_0 agent then chooses to play the action a^i that maximizes the expected value. That is, a ToM_0 agent with zero-order beliefs $b^{(0)}$ chooses to play action

$$t_i^{(0)}(b^{(0)}) = \arg\max_a EV^{(0)}(a; b^{(0)}). \tag{2}$$

Whenever a ToM_0 agent observes the behavior of its opponent, it updates its zero-order beliefs $b^{(0)}$. For example, when ToM_0 agent i observes player j playing action a^j, she updates her beliefs so that

$$b^{(0)}(a) := \begin{cases} (1-\lambda) \cdot b^{(0)}(a) + \lambda & \text{if } a = a^j, \\ (1-\lambda) \cdot b^{(0)}(a) & \text{otherwise.} \end{cases} \tag{3}$$

Note that the ToM_0 strategy is similar to the other-regarding strategy described in Section 3.1. Like the other-regarding strategy, the ToM_0 strategy is drawn to the action that would have won in the previous round. The distinction between the ToM_0 strategy and the heuristic strategies described in the previous subsection is that a ToM_0 agent has an internal state that summarizes all previously observed behavior of the opponent, while the heuristic other-regarding strategy only reacts to the actions in the most recently observed round of play.

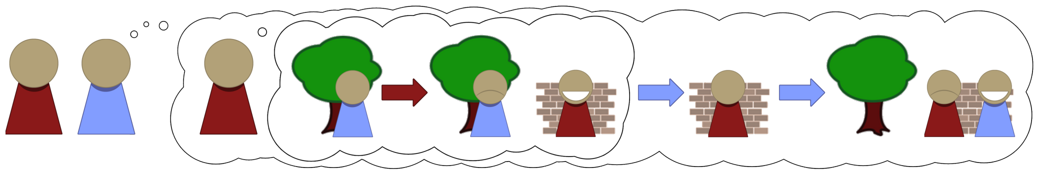

Figure 3: If the blue agent j is a ToM_1 seeker, he puts himself in the position of the hider i to predict where she will hide. If the ToM_1 seeker j usually searches near the tree, the hider i may believe that he is going to search for her near the tree again. If this is what she believes, she would hide behind the wall to avoid the seeker. Therefore, the ToM_1 seeker j concludes, he should seek for the hider behind the wall.

For the purpose of Bayesian RFX-BMS estimation of theory of mind use, we follow Devaine et al. [2] and consider that choices may exhibit small deviations from the optimal decision rule defined by a strategy. We therefore employ the so-called 'softmax' probabilistic policy. That is, the probability that a ToM_0 player i will perform action u is given by

$$P(A^i = u) = s\left(\frac{EV_i^{(0)}(u; b^{(0)})}{\beta}\right) := \frac{\exp\left(EV_i^{(0)}(u; b^{(0)})/\beta\right)}{\exp\left(EV_i^{(0)}(u; b^{(0)})/\beta\right) + \exp\left(EV_i^{(0)}(d; b^{(0)})/\beta\right)}, \tag{4}$$

where β is the exploration temperature, a free parameter that controls the magnitude of behavioral noise.

3.3 First-order theory of mind

In contrast to a ToM_0 agent, a first-order theory of mind (ToM_1) agent can reason about the mental content of its opponent and realize that the opponent has a goal of its own. A ToM_1 agent can place itself in the position of its opponent and calculate what the agent would have done itself. Figure 3 shows an example of this process for a ToM_1 seeker j in the hide-and-seek game. Based on his own previous actions, the ToM_1 seeker j reasons that if he had been in the position of the hider, he would predict that the seeker is going to search behind the tree. To avoid this seeker, the ToM_1 agent would therefore have chosen to hide behind the wall. The ToM_1 seeker j attributes this reasoning process to his opponent and therefore concludes that she is most likely to hide behind the wall. As a result, the ToM_1 seeker decides to search behind the wall.

More in general, to model the beliefs of her opponent, a ToM_1 agent i forms first-order beliefs $b^{(1)}$ that represent what the agent's zero-order beliefs would have been if she had been in the position of her opponent. That is, according to the first-order beliefs $b^{(1)}$ of ToM_1 agent i, if agent i had been in the position of her opponent j, agent i would have believed that the probability that she will perform action a^i is $b^{(1)}(a^i)$. If she had been in the position of her opponent j, she would therefore have chosen to play action $t_j^{(0)}(b^{(1)})$. Using first-order theory of mind, ToM_1 player i predicts that opponent j will do the same.

Based on the observed behavior of her opponent, a ToM_1 player i may come to believe that her first-order beliefs do not accurately predict the behavior of player j. In this case, she may decide to act as if she were a ToM_0 agent instead. This behavior is controlled through the ToM_1 agent's confidence $c_1 \in [0, 1]$ in first-order theory of mind. The higher the confidence c_1, the more the behavior of a ToM_1 agent is determined by first-order beliefs $b^{(1)}$. The expected value $EV_i^{(1)}(a^i)$ that ToM_1 agent i assigns to playing action a^i is

$$EV_i^{(1)}(a^i; b^{(0)}, b^{(1)}, c_1) = c_1 \cdot \pi^i\left(a^i, t_j^{(0)}(b^{(1)})\right) + (1 - c_1) \cdot EV_i^{(0)}(a^i; b^{(0)}). \tag{5}$$

The expected value for ToM_1 player j is constructed analogously. Similar to a ToM_0 agent, the ToM_1 agent chooses to play the action $t_i^{(1)}(b^{(0)}, b^{(1)}, c_1)$ that maximizes its expected value.

After observing the outcome of a game in which the ToM_1 agent i decided to play action a^i and the opponent j played action a^j, a ToM_1 agent i updates her confidence c_1 in first-order theory of mind so that

$$c_1 := \begin{cases} (1-\lambda)\cdot c_1 + \lambda & \text{if } a^j = t_j^{(0)}(b^{(1)}) \\ (1-\lambda)\cdot c_1 & \text{otherwise.} \end{cases} \tag{6}$$

That is, if the agent's first-order theory of mind accurately predicted that opponent j would play a^j, the ToM_1 agent increases her confidence c_1 in first-order theory of mind. Otherwise, the agent lowers her confidence.

Next, the ToM_1 agent i updates her zero-order beliefs as described in Equation (3). In the same way, the ToM_1 agent i updates her first-order beliefs $b^{(1)}$ to increase the belief that she will play action a^i again. Note that the ToM_1 agent i updates her first-order beliefs $b^{(1)}$ using her own learning speed λ. Unlike Devaine's Bayesian agents [2], our *ToM* agents do not attempt to estimate the learning speed λ of their opponent.

Similar to the procedure of the ToM_0 agent, we use a softmax policy with exploration temperature β, so that the probability that a ToM_1 player i will perform action u is given by

$$P(A^i = u) = s\left(EV_i^{(1)}(u; b^{(0)}, b^{(1)}, c_1)/\beta\right), \tag{7}$$

where β is the exploration temperature.

3.4 Higher orders of theory of mind

A kth-order theory of mind (ToM_k) agent considers the possibility that its opponent is a ToM_{k-1} agent in addition to the possibility that the opponent is reasoning at even lower orders of theory of mind. To predict the behavior of its opponent, a ToM_k agent takes the perspective of the opponent and determines what it would have done itself, based on its own kth-order beliefs $b^{(k)}$ and confidence c_k in kth-order theory of mind.

Analogous to the way a ToM_1 agent integrates the predictions of different orders of theory of mind, a ToM_k player i calculates the expected value $EV^{(k)}(a^i)$ of playing actions a^i as

$$\begin{aligned} EV_i^{(k)}(a^i; b^{(0)}, \ldots, b^{(k)}, c_1, \ldots, c_k) = \quad & c_k \cdot \pi^i\left(a^i, t_j^{(k-1)}(b^{(1)}, \ldots, b^{(k)}, 1, 0, \ldots, 0)\right) + \\ & (1-c_k)\cdot EV_i^{(k-1)}(a^i; b^{(0)}, \ldots, b^{(k-1)}, c_1, \ldots, c_{k-1}), \end{aligned} \tag{8}$$

and decides to play the action $t_i^{(k)}(b^{(0)}, \ldots, b^{(k)}, c_1, \ldots, c_k)$ that maximizes this expected value.

Following our previous work [22], our *ToM* agents do not attempt to model their opponent's confidence in theory of mind. Rather, a ToM_k agent i maintains $k+1$ models of her opponent and attempts to find the one that most accurately predicts his behavior.

After observing the outcome of a game in which player i played action a^i and player j played action a^j, the ToM_k agent updates its confidences c_n in nth-order theory of mind according to Equation (6) as well as the corresponding nth-order beliefs $b^{(n)}$. For each even numbered order of theory of mind n, player i updates her beliefs $b^{(n)}$ according to Equation (3) to reflect that she believes it to be more likely that her opponent will play action a^j again. For each odd numbered order of theory of mind m, player i updates her beliefs $b^{(m)}$ to reflect that she predicts her opponent to believe that she is more likely to play action a^i again.

To obtain the likelihood that a ToM_k player i will play action u, we use the softmax policy so that

$$P(A^i = u) = s\left(EV_i^{(k)}(u; b^{(0)}, \ldots, b^{(k)}, c_1, \ldots, c_k)/\beta\right), \tag{9}$$

where β is the exploration temperature.

4 Results

To determine the extent to which theory of mind is used when playing matching pennies, the strategies described in Section 3 are used as the basis for random-effects Bayesian model selection [17]. These strategies

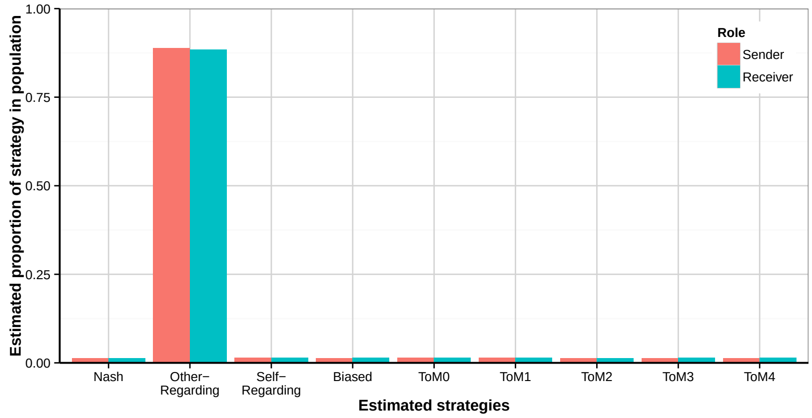

Figure 4: Estimated proportions of strategies used by confederates in the sender-receiver game [16].

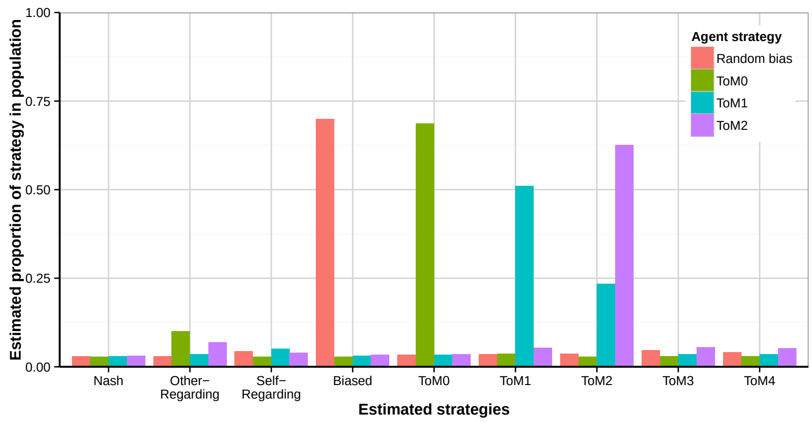

Figure 5: Estimated proportions of strategies used by Devaine's Bayesian agents in the hide-and-seek game and the casino game [2].

include two randomizing strategies (Nash, biased), three behavior-based strategies (other-regarding, self-regarding, ToM_0), as well as four theory of mind strategies (ToM_k, $k \in \{1, 2, 3, 4\}$). Note that only the Nash strategy is parameter-free. The biased, other-regarding, and self-regarding strategies each have a single parameter λ, while the ToM_k strategies ($0 \leq k \leq 4$) have two free parameters: λ and β. The values of these parameters were allowed to vary between subjects, but were assumed to be fixed within subjects.

The experimental data from Devaine et al. [2] and Sher et al. [16] contain the behavioral responses of human participants following an unknown strategy, but also the actions performed by their opponents, who strictly follow a fixed strategy. In Section 4.1, we start by estimating strategy use of these opponents to show that Bayesian RFX-BMS estimation can successfully recover known strategies from behavioral data. In Section 4.2, we estimate the strategies used by human participants in these matching pennies games.

4.1 Validation

To show that Bayesian RFX-BMS estimation can successfully recover player strategies, we apply this method to the behavior of players that follow a known strategy. In the experimental study of Sher et al. [16], children played the sender-receiver game against a confederate who was instructed to follow a fixed other-regarding strategy. After the initial choice, the confederate would always select the action that would have won in the previous round. That is, confederates followed an other-regarding strategy with learning speed $\lambda = 1.0$.

Figure 4 shows the estimated proportions of each of the nine strategies we consider based on the behavioral data of confederates in the sender-receiver game. As the figure shows, Bayesian RFX-BMS estimation successfully recovers the confederate other-regarding strategy. In addition, the average estimated value of the learning speed parameter λ was 0.98. This shows that Bayesian RFX-BMS estimation can successfully recover a player strategy if it is included in the estimation as a population strategy.

The experimental study of Devaine et al. [2] also included players following a known strategy. Both in the hide-and-seek task and in the casino task, participants played against four different computer agents. These agents included a random biased agent that always chose one of the options with a 65% probability and three Bayesian theory of mind agents. Note that although Devaine's Bayesian agents make use of theory of mind, the exact specifications of these agents differ from our *ToM* agent descriptions in Section 3. That is, the strategies used by Devaine's Bayesian agents are not included as population strategies and can therefore not be recovered from the empirical data. This allows us to test whether Bayesian RFX-BMS estimation can accurately classify a strategy as a theory of mind strategy instead of a heuristic strategy, even when the details of the theory of mind strategy used by the player differs from our *ToM* agent definition. In addition, by using different agents to generate behavior and to classify behavior, we can determine to what extent different orders of theory of mind reasoning are consistent among our *ToM* agents and Devaine's Bayesian agents.

Figure 5 shows the results of Bayesian RFX-BMS estimation on the behavioral data of computer agents in the Devaine et al. study, aggregated across the two task settings. The figure shows that the Bayesian RFX-BMS estimation method accurately distinguishes between theory of mind strategies and simpler, heuristic-based strategies. In addtion, the theory of mind abilities of Devaine's Bayesian agents are estimated remarkably well, despite the differences in agent specification. In our Bayesian RFX-BMS estimation results, only Bayesian first-order theory of mind agents are regularly misclassified, and identified as ToM_2 agents. This suggests that Devaine's Bayesian first-order theory of mind agent may be capable of more complex opponent modeling than the ToM_1 agent described in Section 3. Interestingly, Bayesian zero-order and second-order theory of mind agents are rarely misclassified, which suggests a good fit between the two models of theory of mind.

The results in this section show that Bayesian RFX-BMS estimation can accurately recover known strategies used when playing matching pennies, as well as accurately distinguish between theory of mind strategies and heuristic strategies. This suggests that Bayesian RFX-BMS estimation may be useful in determining the extent to which human participants make use of theory of mind when playing this game. In the following subsection, we will consider this issue in detail.

4.2 Participant strategy use

The results of Bayesian RFX-BMS estimation on behavioral data of known strategies suggests that this method can accurately recover player strategies across several variants of the matching pennies game. In this section, we apply Bayesian RFX-BMS estimation to human participant data to determine the extent to which they make use of theory of mind.

Sher et al. [16] let children play a sender-receiver game against a confederate following a strict other-regarding strategy. Note that the best response against this strategy is to follow a self-regarding strategy that alternates between the two possible actions. Figure 6 shows the results of Bayesian RFX-BMS estimation of the strategies used by these children. The figure shows that Bayesian RFX-BMS estimation classifies over 40% of these children as ToM_1 agents, both in the sender role and in the receiver role. However, Figure 6 shows that in the sender role, 24% of the children are classified as using an other-regarding strategy and 17% as using a self-regarding strategy. In the receiver role, 12% of the children are classified as using an other-regarding strategy and 26% as using a self-regarding strategy. That is, the behavior of many of these children is best described as the use of a behavior-based strategy. Note, however, that children are rarely classified as using a randomizing strategy.

Devaine et al. [2] let participants play against Bayesian theory of mind agents in the social setting of a hide-and-seek task and the non-social setting of a casino task. As in the previous subsection, we aggregate data across the two tasks. Unlike Devaine et al., however, we do not aggregate across opponent strategy.

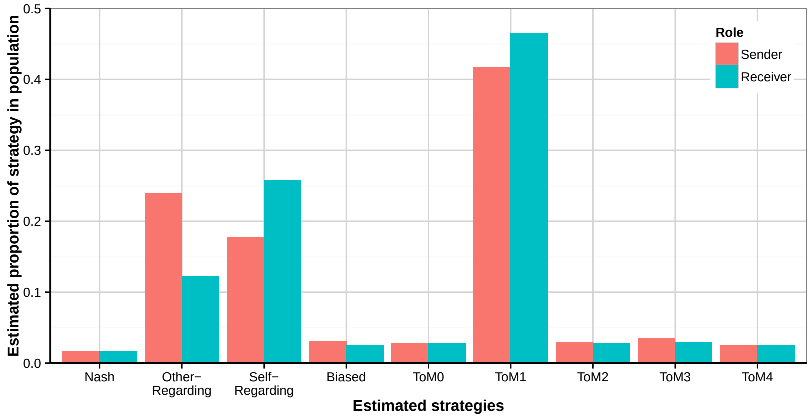

Figure 6: Estimated proportions of strategies used by children playing the sender-receiver game [16].

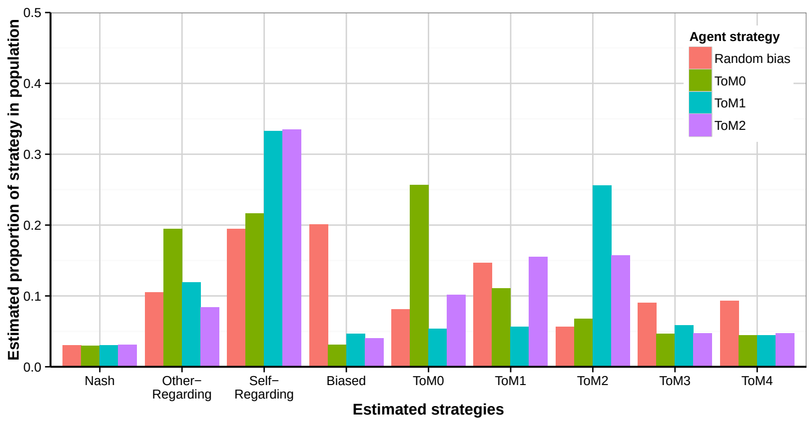

Figure 7: Estimated proportions of strategies used by participants playing the hide-and-seek game and the casino game [2].

That is, we explicitly take into account that participants may adjust their strategy based on the observed behavior of an opponent, while we ignore the effects of task context. Figure 7 shows the results of Bayesian RFX-BMS estimation of the strategies used by the participants. These results show strong variation in the strategy use of participants. Interestingly, participants appear to rarely use randomizing strategies (Nash or biased), with the exception of participants that play against a biased opponent. Note that in this case, the best response indeed is for the participants to follow a biased strategy as well.

Compared to the estimated strategy use of children in Figure 6, adult participants are less often classified as following theory of mind strategies. One notable exception is when participants play against a Bayesian first-order theory of mind agent, where the proportion of participants using a second-order theory of mind strategy is estimated at 25%.

The results in Figure 7 suggest that participants engage in some opponent modeling. Participants that play against a biased agent are more likely to be classified as using a biased strategy themselves, while participants that play against a first-order Bayesian theory of mind agent are more likely to be classified as using second-order theory of mind. However, across the different opponent strategies, many participants are well-described as using a self-regarding strategy. This suggests that a sizable proportion of participants may rely on simple heuristics when playing matching pennies in the context of the hide-and-seek or the casino game, irrespective of the behavior of the opponent.

To summarize, the results of our Bayesian RFX-BMS estimation suggest that there are both child and adult participants that make use of theory of mind strategies while playing matching pennies. However, there also appear to be many participants whose behavior is better described as following a simpler heuristic strategy.

5 Conclusion

Both empirical and simulation studies suggest that players can benefit from reasoning about unobservable mental content of opponents in simple competitive games such as matching pennies or rock-paper-scissors [2–5, 7, 22]. In this paper, we combined computational agents with Bayesian RFX-BMS estimation [17] to determine to what extent human participants actually make use of this so-called 'theory of mind' when playing repeated versions of matching pennies in two different empirical studies.

Sher et al. [16] let children play the sender-receiver game against a human confederate following a fixed strategy, while Devaine et al. [2] let participants play against computational Bayesian theory of mind agents in both a social hide-and-seek task and a non-social casino task. Our results show that Bayesian RFX-BMS estimation can accurately recover the known strategies used by both confederates and computational agents. In addition, our results show that Bayesian RFX-BMS estimation can accurately identify a theory of mind strategy when the model used to estimate theory of mind use differs from the model that generated theory of mind behavior. This suggests that Bayesian RFX-BMS estimation can be used effectively to determine theory of mind use in human participants, even when the the details of their theory of mind strategy are unknown.

Our results suggest that both children and adults engage in theory of mind when playing simple repeated games such as matching pennies. Many children make use of first-order theory of mind in the sender-receiver game, both in the role of sender and in the role of receiver. In addition, adult participants appear to make use of second-order theory of mind when facing a Bayesian first-order theory of mind agent. However, both the child data and the adult data also show that many participants are better described as making use of a simpler behavior-based strategy. In addition, Devaine et al. [2] report that adult participants won, on average, 51% of the games they played. This may indicate that human players do not use their theory of mind abilities primarily to compete with others in one-shot situations such as matching pennies, or may even be unable to profitably apply theory of mind in these game situations.

One reason for the apparent lack of participants' use of theory of mind is that there are many viable simple strategies available in the matching pennies setting. The availability of these simple strategies may make it less appealing for participants to select a more complex theory of mind strategy. In addition, the wide range of strategies may make it more difficult for participants to use theory of mind to accurately predict the behavior of an opponent [cf. 6]. Future research could disentangle the effects of the appeal of using simple strategies from effects of the unpredictability of other players. Settings such as Marble Drop [13], for example, have fewer viable simple strategies than matching pennies. In such settings, less sophisticated players would therefore exhibit more predictable behavior, which may encourage human players to make more use of theory of mind as well.

Alternatively, human participants may be more likely to use their theory of mind in more cooperative settings, or in settings in which both cooperation and competition play a role[1]. Bayesian RFX-BMS estimation could provide more insight in the kind of situations that encourage human participants to make use of higher orders of theory of mind.

References

[1] Camerer, C.F., Ho, T.H., Chong, J.K.: A cognitive hierarchy model of games. The Quarterly Journal of Economics 119(3), 861–898 (2004)

[1]For first results in this direction, see [20, 21].

[2] Devaine, M., Hollard, G., Daunizeau, J.: The social Bayesian brain: Does mentalizing make a difference when we learn? PLoS Computational Biology 10(12), e1003992 (2014)

[3] Devaine, M., Hollard, G., Daunizeau, J.: Theory of mind: Did evolution fool us? PloS ONE 9(2), e87619 (2014)

[4] Franke, M., Galeazzi, P.: On the evolution of choice principles. In: Szymanik, J., Verbrugge, R. (eds.) Proceedings of the Second Workshop Reasoning About Other Minds: Logical and Cognitive Perspectives, co-located with Advances in Modal Logic, Groningen, The Netherlands. CEUR Workshop Proceedings, vol. 1208, pp. 11–15 (2014)

[5] Frey, S., Goldstone, R.L.: Cyclic game dynamics driven by iterated reasoning. PloS ONE 8(2), e56416 (2013)

[6] Ghosh, S., Heifetz, A., Verbrugge, R.: Do players reason by forward induction in dynamic perfect information games? In: Ramanujam, R. (ed.) Proceedings of the 15th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2015). pp. 121–130 (2015)

[7] Goodie, A.S., Doshi, P., Young, D.L.: Levels of theory-of-mind reasoning in competitive games. Journal of Behavioral Decision Making 25(1), 95–108 (2012)

[8] Hedden, T., Zhang, J.: What do you think I think you think?: Strategic reasoning in matrix games. Cognition 85(1), 1–36 (2002)

[9] Heifetz, A.: Game Theory: Interactive Strategies in Economics and Management. Cambridge University Press (2012)

[10] Kinderman, P., Dunbar, R.I., Bentall, R.P.: Theory-of-mind deficits and causal attributions. British Journal of Psychology 89(2), 191–204 (1998)

[11] Meijering, B., van Maanen, L., van Rijn, H., Verbrugge, R.: The facilitative effect of context on second-order social reasoning. In: Ohlsson, S., Catrambone, R. (eds.) Proceedings of the 32nd Annual Conference of the Cognitive Science Society. pp. 1423–1428 (2010)

[12] Meijering, B., van Rijn, H., Taatgen, N.A., Verbrugge, R.: I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In: Carlson, L., Hoelscher, C., Shipley, T.F. (eds.) Proceedings of the 33rd Annual Conference of the Cognitive Science Society. pp. 2486–2491 (2011)

[13] Meijering, B., Taatgen, N.A., van Rijn, H., Verbrugge, R.: Modeling inference of mental states: As simple as possible, as complex as necessary. Interaction Studies 15, 455–477 (2014)

[14] Premack, D., Woodruff, G.: Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences 1(04), 515–526 (1978)

[15] Rapoport, A., Budescu, D.V.: Randomization in individual choice behavior. Psychological Review 104(3), 603–617 (1997)

[16] Sher, I., Koenig, M., Rustichini, A.: Children's strategic theory of mind. Proceedings of the National Academy of Sciences 111(37), 13307–13312 (2014)

[17] Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J.: Bayesian model selection for group studies. Neuroimage 46(4), 1004–1017 (2009)

[18] Stiller, J., Dunbar, R.I.: Perspective-taking and memory capacity predict social network size. Social Networks 29(1), 93–104 (2007)

[19] Wagenaar, W.A.: Generation of random sequences by human subjects: A critical survey of literature. Psychological Bulletin 77(1), 65–72 (1972)

[20] de Weerd, H., Broers, E., Verbrugge, R.: Savvy software agents can encourage the use of second-order theory of mind by negotiators. In: Noelle, D., Dale, R., Warlaumont, A., Yoshimi, J., Matlock, T., Jennings, C., Maglio, P. (eds.) Proceedings of the 37th Annual Conference of the Cognitive Science Society. pp. 542–547 (2015)

[21] de Weerd, H., Verbrugge, R., Verheij, B.: Negotiating with other minds: The role of recursive theory of mind in negotiation with incomplete information. Autonomous Agents and Multi-Agent Systems, In press

[22] de Weerd, H., Verbrugge, R., Verheij, B.: How much does it help to know what she knows you know? An agent-based simulation study. Artificial Intelligence 199-200, 67–92 (2013)

[23] Wright, J.R., Leyton-Brown, K.: Beyond equilibrium: Predicting human behavior in normal-form games. In: Proceedings of the 24th Conference on Artificial Intelligence. pp. 901–907 (2010)