CHAPTER 3

Theory of mind in the Mod game

An agent-based model of strategic reasoning

Abstract

When people engage in social interactions, they often rely on their theory of mind, their ability to reason about unobservable mental content of others such as beliefs, goals, and intentions. This ability allows them to both understand why others behave the way they do as well as predict future behavior. People can also make use of higher-order theory of mind by applying theory of mind recursively, and reason about the way others make use of theory of mind such as in the sentence “Alice believes that Bob does not know about the surprise party”. In this paper, we use agent-based models to describe human behavior in an n-player extension of rock-paper-scissors called the Mod game. In previous work, we have shown how in similar competitive settings, the ability to make use of higher orders of theory of mind can be beneficial. We find that characteristic cyclic behavior in the choices of participants that contradicts equilibrium predictions from classical game theory can be explained through the application of higher orders of theory of mind. Our results suggest that participants engage in higher orders of theory of mind reasoning in repeated play of the Mod game than previously reported in normal-form games and in repeated rock-paper-scissors games.

This chapter has previously been published in Proceedings of the European Conference on Social Intelligence (ECSI-2014).
3.1 Introduction

People often make use of theory of mind (Premack & Woodruff, 1978) to explain and predict the behavior of others. By reasoning explicitly about unobservable mental content such as beliefs, goals, and intentions, people are able to, for example, distinguish accidental from intentional behavior. People can also make use of higher-order theory of mind, by using their own theory of mind ability to reason about the way others may use their theory of mind. Second-order theory of mind allows people to form nested beliefs such as “Alice believes that Bob does not know about the surprise party”, and use these beliefs to understand and predict the behavior of Alice.

Empirical evidence shows that participants can make use of higher-order (i.e. at least second-order) theory of mind in tasks that require explicit reasoning about belief attributions of others (Apperly, 2010; Perner & Wimmer, 1985) as well as in strategic games (Hedden & Zhang, 2002; Meijering et al., 2011, 2010). However, there are limits to the depth of recursive theory of mind reasoning that people use (Keysar et al., 2003; Ohtsubo & Rapoport, 2006). In particular, people are in general unable to explicitly use the infinite recursion needed to reason about common knowledge (Gintis, 2009; Verbrugge, 2009). As a result, people often fail to behave as predicted by equilibrium predictions of classical game theory.

In previous work, we presented an agent-based model of theory of mind (De Weerd, Verbrugge, & Verheij, 2013b) to show that agents benefit from the ability to make use of higher-order theory of mind in certain competitive settings. This model is closely related to hierarchical models of iterated reasoning in behavioral game theory, such as level-\(k\) reasoning (Costa-Gomes, Crawford, & Broseta, 2001; Stahl & Wilson, 1995), cognitive hierarchies (Camerer et al., 2004), quantal response equilibrium (McKelvey & Palfrey, 1995), and noisy introspection models (Goeree & Holt, 2004). In each of these models, the level of sophistication of agents is measured by the maximum number of steps of iterated reasoning the agent is capable of considering. In terms of theory of mind, one step of iterated reasoning approximately corresponds to zero-order theory of mind. Camerer et al. (2004) estimate the distribution of the level of sophistication used by human participants over a range of normal-form games such as the \(p\)-beauty contest and the traveler’s dilemma, and find an average of 1.5 steps of iterated reasoning. Part of the participants were found to use more than two steps of iterated reasoning, which suggests these participants may have been relying on second-order theory of mind. However, only few players were found to be well-described as higher-level agents (Wright & Leyton-Brown, 2010).

Our model of theory of mind agents differs from previous models in that the behavior of our agents changes based on the observed behavior of others. Previous models typically assume that the most basic agent responds optimally under
the assumption that other agents act according to a known non-strategic policy (Camerer et al., 2004; Costa-Gomes et al., 2001; Goeree & Holt, 2004; McKelvey & Palfrey, 1995; Stahl & Wilson, 1995; Yoshida et al., 2008). Instead, our zero-order theory of mind agents attempt to learn the behavior of others in repeated games through heuristics and associative learning. Our agent model has shown that in competitive settings such as repeated rock-paper-scissors games, second-order theory of mind can benefit agents greatly (De Weerd et al., 2013b), while human behavior in normal-form games suggests lower levels of recursive reasoning (Camerer et al., 2004). In the current paper, to determine whether human participants make use of higher-order theory of mind in these competitive settings, we describe human behavior in the Mod game (Frey & Goldstone, 2013), an n-player extension of rock-paper-scissors, using our agent model. By comparing the behavior of theory of mind agents with participant data from the Mod game, we can determine to what extent higher orders of theory of mind can account for the observed patterns of human behavior.

In Section 3.2, we present a detailed description of the Mod game, as well as the way human participants play the game. Section 3.3 shows how theory of mind agents are implemented in this setting and how these agents play the Mod game. We simulate interactions between these theory of mind agents and compare the results to the behavior of human participants. The results of these experiments are outlined in Section 3.4. Finally, Section 3.5 provides a discussion of the results as well as directions for future research.

### 3.2 Mod game

The Mod game is an n-player generalization of rock-paper-scissors, introduced by Frey & Goldstone (2013) as a way to reveal patterns in individual iterated reasoning strategies. In the Mod game, n participants simultaneously choose a number in the range \{1, \ldots, m\}, for both n and m greater than one. Participants gain one point for every other participant that has chosen the number that is exactly one lower than their own choice. For example, a participant that has chosen the number 4 gains a point for every participant that has chosen number 3. The only exception to this rule is that participants that have chosen number 1 gain one point for every participant that has chosen number m. The Mod game therefore has a structure similar to rock-paper-scissors, in which each action is dominated by some other action. In particular, the Mod game is a non-zero-sum version of rock-paper-scissors for n = 2 and m = 3.

The Mod game has a mixed-strategy Nash equilibrium in which each action is chosen with equal probability. When all players play according to this strategy, none of the players has an incentive to change his or her strategy. However, if a player deviates from the randomizing strategy, other players can take advantage
3.2. Mod game

Figure 3.1: Histograms over 24 choices, rates, and accelerations of human behavior in the Mod game (red curve, reconstructed from Frey, 2013). The blue curve shows the expected results from random behavior.

from this regularity by switching their strategy as well. Interestingly, human participants are generally poor at generating random sequences (Herbranson & Schroeder, 2010; Rapoport & Budescu, 1997; Wagenaar, 1972). This suggests that in groups of people, individuals can increase their score if they deviate from playing the Nash equilibrium strategy.

Indeed, participant behavior in repeated Mod games deviates from the Nash equilibrium, as shown in Figure 3.1 (reconstructed from Frey, 2013). The figure shows the aggregate participant data (red curve) and the idealized randomizing behavior (blue curve) over 100 rounds of play and \( m = 24 \). Participant choices appear to be approximately random, with a slight bias towards 24. However, participant rates, defined as the difference in choice between two subsequent rounds, shows a clear deviation from the Nash equilibrium in Figure 3.1. Participants are less likely to select numbers that are 7 to 21 ahead of their previous choice. Instead, they are most likely to choose a number that is 0 to 4 higher than their previous choice. Participant acceleration, which is defined as the change in participant rate, shows a similar effect. Figure 3.1 shows that participants tend to vary little in their rate. A participant who chose a number in the last round that was 2 higher than the number in the round before that is mostly likely to choose the number that is 2 higher than his choice in the previous round. However, Figure 3.1 shows that participants do vary their acceleration by a small amount.

Interestingly, this effect is not due to participants’ poor performance on choosing random actions. When participants are given the option to let the computer select a randomly generated action, this option is used little (Frey, 2013). This suggests that participants choose their actions based on their predictions of the behavior of others rather than believing that the behavior of others is unpredictable.
3.3 Theory of mind agents in the Mod game

In this section, we describe theory of mind agents that are able to play the Mid game outlined in Section 3.2. These agents are inspired by the theory of mind agents that we introduced in De Weerd et al. (2013b) to investigate the effectiveness of theory of mind in competitive settings. They engage in simulation-theory of mind (Davies, 1994; Nichols & Stich, 2003; Hurley, 2008) by taking the perspective of their opponents. An agent determines what he would do himself if he were facing the situation of his opponent, and attributes this thought process to his opponent to predict her behavior. Each additional order of theory of mind allows the agent to generate an additional hypothesis about the way an opponent is playing the game. The task of a theory of mind agent is then to determine which hypothesis most accurately predicts the behavior of his opponent.

The following subsections describe how agents of different orders of theory of mind play the $n$-player Mod Game (Frey & Goldstone, 2013).

3.3.1 Zero-order theory of mind

A zero-order theory of mind ($ToM_0$) agent has no theory of mind at all, and is therefore unable to attribute mental content to others. In particular, a $ToM_0$ agent cannot form the belief that his opponents are trying to obtain a high score. Instead, the $ToM_0$ agent forms zero-order beliefs about the collective actions of the agents playing the game. For each number, the $ToM_0$ agent specifies what he believes to be the likelihood that most players will select to play that number. Given these beliefs, the $ToM_0$ agent can select the number that he expects to maximize his score. For example, if a $ToM_0$ agent strongly believes that number 4 will be selected by most players, the agent should choose to play number 5 himself.

After every round, the $ToM_0$ agent updates his zero-order beliefs to reflect the actual outcome, such that the agent’s new beliefs are constructed through a linear combination of his original beliefs and the newly observed situation. An agent-specific learning speed $\lambda \in [0, 1]$ determines the relative influence of the observation on the agent’s beliefs. A $ToM_0$ agent with zero learning speed ($\lambda = 0$) does not update his beliefs at all. Such an agent selects the same action in every round. A $ToM_1$ agent with the maximal learning speed ($\lambda = 1$), on the other hand, completely replaces his zero-order beliefs after each observation, and forgets all information obtained from previous rounds. Such an agent considers the observed actions of the last round as the best predictor for the future.

The $ToM_0$ agent we describe here holds a simple model of agent behavior. Although agents could model the behavior of each individual opponent separately,
our $ToM_0$ agent models the collective behavior of others instead$^1$. The $ToM_0$ agents also do not have any explicit memory. Although information obtained in the past is reflected in the agent’s zero-order beliefs, the agent does not have an explicit representation of the past.

### 3.3.2 First-order theory of mind

Unlike the $ToM_0$ agent, a first-order theory of mind ($ToM_1$) agent reasons about the other’s goals and therefore believes that his opponents may be trying to maximize their score as well. To predict the behavior of his opponents, the $ToM_1$ agent attributes his own thought process to his opponents. Like the $ToM_0$ agent, the $ToM_1$ agent models a modal opponent rather than having a separate mental model for each individual opponent.

Following De Weerd et al. (2013b), the $ToM_1$ agent does not attempt to model the learning speed $\lambda$ for his first-order model of opponent behavior. Rather, the $ToM_1$ agent assumes that the modal opponent has the same learning speed as he has himself. This means that after a sufficient number of rounds, the $ToM_1$ agent predicts that most of the his opponents will play the action suggested by the agent’s zero-order theory of mind. For example, for a $ToM_1$ agent that has strong zero-order beliefs that number 4 will be selected by most players, the agent’s zero-order response would be to play number 5. Using first-order theory of mind, the $ToM_1$ agent attributes this though process to his opponents, and predicts that most of them will play number 5. The $ToM_1$ agent’s first-order response would therefore be to play number 6.

Although the $ToM_1$ agent models his opponents as being able to use zero-order theory of mind, agents in our setup do not know the extent of the abilities of their opponents for certain. Rather, a $ToM_1$ agent has two models of opponent behavior, one based on zero-order theory of mind and one on first-order theory of mind. Through repeated interaction, a $ToM_1$ agent learns which of his models best describes the behavior of his opponents. Based on this information, a $ToM_1$ agent may therefore choose to play as if he were a $ToM_0$ agent, and ignore the predictions of his first-order theory of mind.

### 3.3.3 Higher orders of theory of mind

For each additional order of theory of mind $k$, an agent generates an additional prediction of opponent behavior by attributing his own $(k-1)$st-order theory of mind thought process to his opponents. For example, a $ToM_2$ agent models the

---

$^1$Empirical evidence from weak-link coordination games suggests that participants only consider part of the data (Devetag, 2005; Van Huyck, Battalio, & Beil, 1990). However, evidence to the contrary also exists (see for example Berninghaus & Ehrhart, 2001; Brandts & Cooper, 2006).
The agent model shows that in competitive settings such as repeated rock-paper-scissors games, second-order theory of mind can benefit agents greatly, while the additional benefit for fourth-order and even higher orders of theory of mind is limited (De Weerd et al., 2013b). In this paper, we therefore do not consider agents that are capable of theory of mind orders beyond the third.

3.4 Simulation results

We simulate interactions of groups of the theory of mind agents of Section 3.3 in the Mod Game described in Section 3.2. To recreate the setting in Frey (2013); Frey & Goldstone (2013), agents repeatedly played the Mod Game for 100 rounds. In each condition, all agents had the same order of theory of mind, and were assigned randomized learning speeds. In Frey (2013), participants had the option to let the computer generate a random choice rather than to choose themselves, which was used in 9% of all choices. To simulate this, each action was assigned a 9% probability of being replaced with a randomly chosen action. That is, each agent had a 9% probability of selecting a random action rather than the one he believes to be the best action.

Figure 3.2 and Figure 3.3 show histograms of the simulations that we have run with groups of agents that share the same ability to make use of theory of mind, as well as data from human participants reported in Frey (2013). In each panel, the red curve shows the data obtained from participants or agent simulations, while the blue curve shows the histogram corresponding to the unique Nash equilibrium.

The first panel of each triple shows a histogram over the possible choices, aggregated over 100 rounds of play. Even though the theory of mind agents only have a 9% probability of choosing randomly, their choices superficially appear to resemble a random distribution, irrespective of the order of theory of mind of the agents.

The second panel of each triple shows the rate at which agents and participants change their choice. That is, these panels show the first difference in the choice of the participants and agents. The third panel shows the acceleration of agents and participants, which is the first difference in the rate of change of the choice of agents and participants.

Figure 3.2 shows that although $ToM_0$ and $ToM_1$ agents superficially appear to make random choices, these agents select their choices in a predictable pattern. Figure 3.2b shows that $ToM_0$ agents typically select the number that is one higher
3.4. Simulation results

(a) Participant data (reconstructed from Frey & Goldstone (2013))

(b) Group of five ToM$_0$ agents

(c) Group of five ToM$_1$ agents

Figure 3.2: Histograms over 24 choices, rates, and accelerations. In each graph, the blue curve shows the expected results from random behavior, while the red curve shows the agent or participant behavior.
3.4. Simulation results

Figure 3.3: Histograms over 24 choices, rates, and accelerations. In each graph, the blue curve shows the expected results from random behavior, while the red curve shows the agent or participant behavior.

than the number that was chosen the most often in the previous round. That is, they select the number that would have won in the previous game. Since each $ToM_0$ agent acts the same way, $ToM_0$ agents typically have a rate of 1 and an acceleration of 0.

Figure 3.2c shows that $ToM_1$ agents display more variation in their behavior. A $ToM_1$ agent typically selects the number that is either one or two higher than the number that was chosen most often in the previous round. The third panel of Figure 3.2c shows that, unlike $ToM_0$ agents, $ToM_1$ agents do not have zero acceleration. Rather, $ToM_1$ agents switch between choosing the number that would have won in the previous game and the number that is one higher. Although the rate data of $ToM_1$ agents fit participant data in Figure 3.2a better than
Chapter 3

3.5. Discussion

ToM₀ agent data, acceleration data from ToM₁ agents differ from participant data. Whereas participants are most likely to keep their rate constant, ToM₁ agents are less likely to do so. Instead, ToM₁ agents are more likely to alternate between increasing and decreasing their rate.

For increasingly higher orders of theory of mind, agent behavior shows increasingly more variation. Although the simplified agent architecture does not reach the variability seen in participant data, results from the higher-order theory of mind agent simulations show some interesting similarities to participant data. Similar to participants, higher-order theory of mind agents typically select numbers that are a little higher than the one they chose in the previous round. More precisely, ToM₂ agents tend to select a number that is up to 3 higher than their previous choice, while ToM₃ agents are also likely to select a number that is 4 higher than their previous choice. Note that this is true even though agents cannot remember their previous choices. Also, acceleration data from participants more closely resembles the acceleration data from higher-order theory of mind agents than that of ToM₀ and ToM₁ agents. Like participants, ToM₂ agents and ToM₃ agents display variation in their acceleration, but they are most likely to keep their rate constant. Interestingly, the higher-order agents also show the average negative acceleration reported in Frey (2013).

Our results show that the behavior of theory of mind agents displays some interesting similarities to the behavior of human participants in repeated Mod games. Moreover, human behavior in the Mod game is closer to the behavior of agents that make use of higher-order theory of mind than that of agents that are limited to zero-order or first-order theory of mind.

3.5 Discussion

Experimental studies show that people make use of higher-order theory of mind. In previous work, we have shown that in certain settings, agents can benefit from using higher orders of theory of mind. In cooperative settings, for example, first-order and second-order theory of mind can help to establish cooperation faster (De Weerd, Verbrugge, & Verheij, 2015) even when a cooperative solution can be maintained without the use of theory of mind, while higher-order theory of mind can also provide an agent with a competitive advantage over others (De Weerd et al., 2013b). In this paper, we have compared these theory of mind agents with human behavior in repeated play of the Mod game (Frey & Goldstone, 2013).

The Mod game has a mixed-strategy Nash equilibrium in which each action is played with equal probability. Contrary to predictions of classical game theory, participants do not appear to randomize their decision in every game. Rather, participant choices show cycles of varying speed (Frey & Goldstone, 2013). Interestingly, our simulations with groups of theory of mind agents show that this
kind of behavior is a closer match to the behavior of a group of agents capable of at least second-order theory of mind than to the behavior of a group of agents that is more limited in their theory of mind abilities.

In previous work, the average depth of reasoning used by humans over a range of games was determined to be 1.5 steps (Camerer et al., 2004), which corresponds approximately to first-order theory of mind. A recent study into human behavior in repeated rock-paper-scissors games suggests that people similarly reason at low orders of theory of mind (Batzilis, Jaffe, Levitt, List, & Picel, 2014). However, whereas participants in previous studies played unrepeated games or games that were repeated three times, participants played more repetitions in the Mod game. Our results suggest that over repeated plays of competitive games such as the Mod game, participants may increase their depth of reasoning and make use of higher orders of theory of mind.

For competitive settings such as the Mod game, our agent model suggests that reasoning at higher orders of theory of mind can benefit agents up to a certain point (De Weerd et al., 2013b). In particular, we found no advantage for the use of fourth-order theory of mind. Simulations in mixed-motive settings, in which both cooperative and competitive goals play a role, suggest that theory of mind also helps to stabilize mutually beneficial interactions (De Weerd, Verbrugge, & Verheij, 2013a). In contrast to strictly competitive settings, reasoning using fourth-order theory of mind may be beneficial in negotiations (De Weerd, Verbrugge, & Verheij, 2014a). It would be interesting to investigate whether human participants are capable of taking advantage of this benefit.

Acknowledgments

This work was supported by the Netherlands Organisation for Scientific Research (NWO) Vici grant NWO 277-80-001, awarded to Rineke Verbrugge for the project ‘Cognitive systems in interaction: Logical and computational models of higher-order social cognition’.
Part II

Testing the
Vygotskian intelligence hypothesis