If you know what I mean

de Weerd, Hermanes Albertus

CHAPTER 8

Related work

In the previous chapters, we have described an agent model for theory of mind reasoning to determine to what extent different types of scenarios may have contributed to the emergence of higher-order theory of mind in humans. However, our approach is not the only way to model theory of mind. There are similar recursive and hierarchical models of theory of mind reasoning, which have been used successfully to model the behavior of human participants. In Section 8.1, we compare our theory of mind agents to related work, and outline how our approach is similar, and in what way our theory of mind agents differ from existing models of recursive reasoning. In Section 8.2, we discuss findings of agent-based models and empirical results on human theory of mind behavior.

8.1 Models of recursive reasoning

In behavioral economics, recursive modeling of the behavior of others has been successfully modeled through iterated best-response models such as level-n theory (Stahl & Wilson, 1995; Bacharach & Stahl, 2000; Costa-Gomes et al., 2001), cognitive hierarchies (Camerer et al., 2004), quantal response equilibria (McKelvey & Palfrey, 1995), and noisy introspection models (Goeree & Holt, 2004). Similar to theory of mind reasoning, in these models, an agent's level of sophistication is measured by the maximum number of steps of iterated reasoning the agent is capable of considering.

The goal of these models of iterated reasoning in behavioral economics is to describe the level of sophistication of human participants in one-shot non-repeated games, such as the p-beauty contest game. In the p-beauty contest game, participants are asked to simultaneously select a number between 0 and 100. Every participant that selects the number closest to two-thirds of the average of the participants' choices wins the game. Which number will be the winning number therefore depends on the behavior of others. The game derives its name from the Keynesian beauty contest (Keynes, 1936, Chapter 12), in which rational agents are asked to select the six prettiest faces among 100 photographs. Those agents that pick the most popular faces win the game. Keynes argues that although the naive strategy is to pick the faces that the agent considers to be the prettiest, a rational agent can increase its chances of winning by selecting the six faces based on the most popular opinion of beauty.

Level-n theory (Stahl & Wilson, 1995) formalizes this idea. A naive level-0 agent does not reason strategically at all. Instead, the level-0 agent randomly chooses a number between 0 and 100, similar to naive agents that select what they consider to be the prettiest faces in the Keynesian beauty contest. In contrast, a level-1 agent believes that all other participants are level-0 agents, and selects the number that is a best response to that belief. According to the level-1 agent, therefore, the winning number will be close to 33 (two-thirds of 50, the average choice of level-0 agents). A level-2 agent performs two steps of iterated reasoning, and therefore believes that all other participants are level-1 agents. As a result, the level-2 agent believes that the winning number will be close to 22 (two-thirds of 33).

The cognitive hierarchy model (Camerer et al., 2004) allows more sophisticated agents that have beliefs about the relative proportions of lower-level agents. In a p-beauty game, for example, cognitive hierarchy models predict that a level-2 agent believes that there will be level-0 individuals that choose randomly and level-1 individuals that will select a number close to 33. However, such a level-2 agent does not suspect that there may be agents that are even more sophisticated in their reasoning process. As a result, the level-2 agent in cognitive hierarchies will select a number between 22 and 33, depending on its beliefs about the relative proportion of level-0 and level-1 agents.

Investigations into the emergence of theory of mind using level-n agents show that higher levels of sophistication do not automatically result in an evolutionary advantage. Stahl (1993) argues that in an evolutionary sense, being right is more important than being smart. For example, if the optimal behavior of an agent is not dependent on the behavior of others, level-0 agents would eventually exhibit this optimal behavior through the process of natural selection. However, Mohlin (2012) shows that certain symmetric two-player normal form games support the evolution of a population in which both lower-level agents as well as higher-level agents coexist.

Level-n agents and cognitive hierarchy models are aimed at modeling recursive reasoning behavior in one-shot non-repeated games and therefore do not allow for learning. Instead, level-n agents make predictions of the initial choices of participants based on their assumptions about the level of sophistication of other individuals. In contrast, our theory of mind agent model focuses on how agents change their behavior in response to the interaction with other agents. The model of iterated reasoning in behavioral economics and the agent model we have outlined in this thesis can therefore be considered to be complementary models, where the former describes the initial behavior of participants, while the latter attempts to capture the way participants change their behavior over time.

Since our agent model of theory of mind reasoning accounts for learning across repeated games, our approach is more similar to models such as recursive opponent modeling (Gmytrasiewicz & Durfee, 1995; Gmytrasiewicz et al., 1998), interactive POMDPs (I-POMDPs; Gmytrasiewicz & Doshi, 2005), and game theory of mind (Yoshida et al., 2008). Similar to our approach, I-POMDP agents can adjust their level of recursive reasoning, but cannot observe the level of sophistication of other agents directly. Instead, agents infer the order of theory of mind at which other agents reason by matching observed behavior of others to the behavior predicted by the application of theory of mind. This means that for both I-POMDP agents and our theory of mind agents, a ToM_k agent can reason about other agents that make use of orders of theory of mind up to and including $(k-1)$st-order theory of mind and adjust their own behavior accordingly. This means that if a ToM_4 agent believes that some individual is a ToM_1 agent, he may decide to behave as if he were a ToM_2 agent. When agents mutually engage in modeling the order of theory of mind at which the other is reasoning, this may influence the effectiveness of higher orders of theory of mind. Through simulation, these effects can be taken into account.

Table 8.1 shows an overview of the four agent models with the strongest relation to our theory of mind agent. For each of these four models, the table shows how the behavior of ToM_0 agents is determined and how higher-order theory of mind agents learn the level of sophistication of others.

Our approach differs from previous work in that the behavior of our agents continues to change based on the observed behavior of others. Previous models of theory of mind typically assume that the most basic agent responds optimally under the assumption that the behavior of other agents can be considered to be random noise, or that other agents act according a known non-strategic policy (Stahl & Wilson, 1995; McKelvey & Palfrey, 1995; Costa-Gomes et al., 2001; Camerer et al., 2004; Goeree & Holt, 2004; Yoshida et al., 2008; Mohlin, 2012). Instead, our zero-order theory of mind agents attempt to learn the behavior of others through associative learning and continue to learn when the behavior of others begins is inconsistent with previously held beliefs. In addition, a theory of mind agent believes that other individuals may vary their order of theory of mind reasoning over the course of repeated games.

One of the consequences of the adaptivity of the zero-order theory of mind agent is that the exact behavior of a ToM_0 agent depends on the behavior of other individuals. A ToM_0 agent that is surrounded by other ToM_0 agents may behave differently than one that is surrounded by ToM_1 agents. By observing the behavior of more sophisticated theory of mind agents, a zero-order theory of mind agent may learn to behave as if it were using theory of mind, without the need of engaging in theory of mind reasoning itself. For example, the ToM_0 agents in the one-shot negotiation setting of Chapter 5 made offers that were more

Agent model	Behavior ToM_0 agent	ToM_k opponent modeling
Level-n theory	Fixed strategy	Opponent is assumed to be a ToM_{k-1} agent
Cognitive hierarchy	Fixed strategy	Known relative proportions of agents of lower orders
Game Theory of Mind	Fixed policy	Bayesian learning
Interactive POMDP	POMDP	Bayesian learning

Table 8.1: Comparison summary of models of recursive reasoning from the literature.

generous towards their trading partner when they competed with a ToM_1 agent compared to when the competitor was restricted to zero-order theory of mind reasoning. This kind of learning does not happen when ToM_0 agents assume that the behavior of other individuals can be considered to be random noise, or if ToM_0 agents assume other individuals act according to a known strategy.

The adaptivity of lower-order theory of mind agents can also benefit higher-order theory of mind agents. For example, theory of mind agents can speed up the process of reaching equilibrium behavior in a cooperative setting by predicting how lower-order agents react to observed outcomes (see Chapter 4). The adaptive nature of our theory of mind agents can also prevent exploitation by higher-order theory of mind agents. For example, the additional advantage for third-order theory of mind in the competitive settings of Chapter 2 is limited because the ToM_2 agent continuously switches between zero-order, first-order, and second-order theory of mind reasoning.

Since we started our investigation into the emergence of higher-order theory of mind using agent-based models (Verbrugge, 2009; De Weerd & Verheij, 2011), similar questions have been addressed in related work. Franke & Galeazzi (2014) compare the evolutionary success of level-n agents that maximize their expected utility to that of level-n agents that minimize regret, and find that regret minimizing agents can outperform utility maximizing agents. In addition, Franke & Galeazzi find that level-n agents of increasingly higher levels continue to obtain an advantage over level-$(n-1)$ agents. Interestingly, their results also show that populations that only contain high-level reasoners can sometimes be invaded by level-0 agents. This means that under certain circumstances, populations that contain both low-level and high-level agents can be evolutionarily stable.

Pynadath et al. (2013) show how theory of mind abilities can be realized in *Sigma* (Σ), an integrated computational model of intelligent behavior that is grounded in a cognitive architecture (Langley, Laird, & Rogers, 2009). Pynadath et al. show how theory of mind can be modeled both as a fast, automatic process and a slow, deliberative process in Sigma, or, alternatively, both as a System 1 and a System 2 process (Kahneman, 2011). Through iterated reasoning, the Sigma model of theory of mind reasoning is able to uncover Nash equilibria in fully observable two-player games. Following up on this work, Pynadath, Rosenbloom, & Marsella (2014) show that Sigma can also be used to capture theory of mind reasoning in a negotiation setting where agents do not know the goal of their trading partner. In this setting, Pynadath et al. find that the ability to make use of theory of mind is only marginally beneficial when the trading partner is using a stationary policy.

Recently, Devaine et al. (2014b) investigated the effectiveness of higher-order theory of mind using a model of meta-Bayesian agents that is closely related to our agent model. Using replicator dynamics, Devaine et al. determine whether Bayesian agents of a lower order of theory of mind can survive when faced with more sophisticated agents in both a competitive setting and a cooperative setting. In their competitive hide-and-seek setting, their Bayesian theory of mind agents benefit from the ability to make use of increasingly higher orders of theory of mind. In this setting, only agents using the highest order of theory of mind survive in the population. In a cooperative setting, on the other hand, reasoning at higher orders of theory of mind is not always beneficial. Devaine and colleagues find that in the battle of the sexes setting, the population reaches an evolutionarily stable state when two-thirds of the population consists of second-order theory of mind agents while the remaining one-third of the population consists of first-order theory of mind agents.

8.2 Models of human behavior

In the previous section, we compared our agent model of theory of mind to several alternative models of recursive reasoning. In this section, we discuss how these models relate to empirical results on human theory of mind. Note that the main goal of our agent model is to provide insight into the emergence of higher-order theory of mind. As such, our agent model is not meant to study recursive reasoning about the knowledge of others from a prescriptive perspective using formal methods such as (dynamic) epistemic logic (Fagin et al., 1995; Van Ditmarsch et al., 2007), or to be a simulation tool that describes the way human participants make use of theory of mind, such as PsychSim (Pynadath & Marsella, 2005).

Level-n theory and cognitive hierarchy models have been successfully used to describe participant behavior. Over a range of one-shot non-repeated games including the p-beauty contest, participants were estimated to be level 1.5 agents on average (Camerer et al., 2004; Costa-Gomes & Crawford, 2006). In terms of theory of mind reasoning, a zero-order theory of mind agent can be considered to perform roughly one step of iterated reasoning. That is, in these one-shot non-repeated games, the vast majority of participants reasoned at zero-order or first-order theory of mind. Although part of the participants were found to use more than two steps of iterated reasoning, only few players were found to be well-described as higher-level agents (Wright & Leyton-Brown, 2010).

Interestingly, participant behavior in repeated games is more consistent with the behavior of higher-order theory of mind agents than with the behavior of lower-order theory of mind agents (Yoshida et al., 2008; Frey & Goldstone, 2013; Frey, 2013; Devaine et al., 2014b, also see Chapter 3 and Chapter 7 of this thesis). Compared to the empirical results of participant behavior in non-repeated games, participants therefore achieve higher orders of theory of mind reasoning over repeated games than their initial choices suggest. Indeed, empirical results show that participants initially reason at low orders of theory of mind, but can increase their order of theory of mind reasoning over repeated games (Hedden & Zhang, 2002; Zhang et al., 2012; Goodie et al., 2012; Meijering et al., 2010, 2011, 2014; Devaine et al., 2014a).

Yoshida et al. (2008) use an agent-based model to show evidence of higher-order theory of mind reasoning of participants in a cooperative game. In their sequential game variation of the stag hunt game, Yoshida et al. find that participant behavior is most consistent with that of level-5 agents. This would therefore suggest that participants achieve particularly high levels of theory of mind reasoning in complex cooperative games. In contrast to these findings, empirical results also show that participants reach higher orders of theory of mind reasoning when they engage in purely competitive and relatively simple games (Goodie et al., 2012). Although these results appear to be contradictory, the common factor in these settings may be the saliency of the goals of other players. Recent research has shown that participants who are informed that they are in competition with another player reach higher orders of theory of mind reasoning than naive participants (Devaine et al., 2014a). In these competitive games, the goal of a competing player is especially salient since it is exactly the opposite of the participant's own goal. In the cooperative stag hunt, on the other hand, it may be easier to reason about the goal of the other player because the participant has the same goal.

In this chapter, we have discussed models of recursive reasoning that are related to our theory of mind agents and presented simulation and empirical results from this related work. In Chapter 9, we compare these findings to the results of our agent-based simulations.