Reasoning about self and others

Meijering, Ben

Chapter 6

Modeling inference of mental states: As simple as possible, as complex as necessary

Abstract

Behavior oftentimes allows for many possible interpretations in terms of mental states, such as goals, beliefs, desires, and intentions. Reasoning about the relation between behavior and mental states is therefore considered to be an effortful process. We argue that people use simple strategies and thus expend less effort as a way of dealing with limited cognitive resources. To test this hypothesis, we developed a computational cognitive model, which was able to simulate previous empirical findings: People start with simple strategies first, and only start revising their strategies when necessary. The model could simulate these findings by means of an interaction between factual knowledge and problem solving skills. At first, the model only considers its own goal, the most basic problem solving skill. Later, the model learns to attribute its problem solving skills to the other player, which only happens if its successes – stored as factual knowledge in declarative memory – do not increase anymore. The model was validated by means of a comparison with findings of a developmental study. This comparison showed that children use the same simple strategies that the model used. To conclude, the model was able to simulate two empirical findings: (1) People try to use simple strategies to infer mental states of others, and (2) they are able to improve such inference by attributing their own strategies to the other player.

This chapter was submitted to a journal and is currently under revision.

Introduction

In social interactions, we try to understand others' behavior by reasoning about their goals, intentions, beliefs, and other mental states. Reasoning about mental states requires a so-called *theory of mind*, abbreviated ToM (Baron-Cohen, Leslie, & Frith, 1985; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). ToM has been implemented in computational cognitive models before (Hiatt & Trafton, 2010; Van Maanen & Verbrugge, 2010). However, these models either simulated one specific instance of ToM (Hiatt & Trafton, 2010) or attributed too much rationality to human reasoning (Van Maanen & Verbrugge, 2010). Here, we present a model that simulates application of various ToM strategies, ranging from simple strategies to full-blown recursive ToM. It is based on previous empirical results (Meijering, Van Maanen, Van Rijn, & Verbrugge, 2010; Meijering, Van Rijn, Taatgen, & Verbrugge, 2011) and is validated by means of a re-analysis of a previous developmental study by Flobbe et al. (2008). The model can explain why people use strategies that are relatively simple, while still being successful at inferring mental states of others.

Many studies have shown that people cannot always account for another's mental states in order to predict their behavior, particularly in the context of two-player sequential games (e.g., Flobbe et al., 2008; Hedden & Zhang, 2002; Raijmakers, Mandell, Van Es, & Counihan, 2013; Zhang, Hedden, & Chia, 2012). Sequential games require reasoning about complex mental states, because Player 1 has to reason about Player 2's subsequent decision, which in turn is based on Player 1's subsequent decision (Figure 6.1). Typically, performance is suboptimal and that is probably because players do not have a correct model of the other player's mental states (Johnson-Laird, 1983). By means of hypothesis testing, they may try to figure out which model works best in predicting the other player's behavior (Gopnik & Wellman, 1992; Wellman et al., 2001). However, a particular action or behavior can have many possible mental state interpretations (Baker, Saxe, & Tenenbaum, 2009), and testing all these interpretations strains our cognitive resources.

To alleviate cognitive demands, people generally start testing simple models or strategies that have been proven successful before (Todd & Gigerenzer, 2000). Because application of ToM and especially recursive ToM is an effortful process (Keysar, Lin, & Barr, 2003; Lin, Keysar, & Epley, 2010; Qureshi, Apperly, & Samson, 2010), reasoning about mental states probably also comprises the use of simple strategies. So where do these strategies come from? We hypothesize that they are a legacy of our childhood years. Raijmakers et al.'s (2013) findings corroborate this claim, as the children in their study consistently used strategies that were not fit to deal with the logical structure of the games presented to them. The strategies sometimes did yield the best possible outcome, however, which may be an explanation for why they still exist in adult reasoning: Simple strategies do not exhaust cognitive resources and are appropriate in a wide range of circumstances. Indeed, our computational cognitive model will show that the presence of simple strategies depends on the proportion of games in which they yield an optimal outcome.

In this study, we present a computational cognitive model that simulates inference of mental states in sequential games. The model initially uses a simple strategy that ignores many task aspects. However, if the model's strategy does not work, it learns to acknowledge that the other player has a role in its outcome. The model will therefore start attributing its own strategy to the other player. We will show that this process can account for the differential learning effects in Meijering et al.'s study (2011; also see Chapter 2 in this dissertation), in which participants adopted distinct strategies based on the training regimen that was administered to them. To validate the model, the developmental study of Flobbe et al. (2008) was re-analyzed, searching for patterns that are indicative of the use of simple strategies in children.

Before we explain the model, we will first explain the empirical findings on which it is based.

Empirical findings

Meijering et al. (2011) studied second-order ToM reasoning in two-player sequential games. Take the game in Figure 6.1 as an example game: Each end node contains a pair of payoffs, left-side payoffs belonging to Player 1 and right-side payoffs belonging to Player 2. The end node in which a game is stopped determines the payoff each player obtains in that particular game. Each player's goal is to obtain his or her greatest attainable payoff. As a player's outcome depends on the other player's decision, both players have to reason about one another's mental states. Participants are always assigned to the role of Player 1, and decide at the first decision point whether to stop the game at A or to continue to the next decision point, which is Player 2's decision between his payoff in B and his payoff in either C or D, which in turn depends on Player 1's decision between Player 1's payoffs in C and D. Thus, before making a decision at the first decision point, participants have to reason about Player 2, who in turn has to reason about Player 1's subsequent decision. In other words, participants have to apply second-order ToM when making a decision.

Meijering et al.'s study was based on the findings of Hedden and Zhang (2002; 2012) and Flobbe et al. (2008). Flobbe et al. had raised some concerns about Hedden and Zhang's

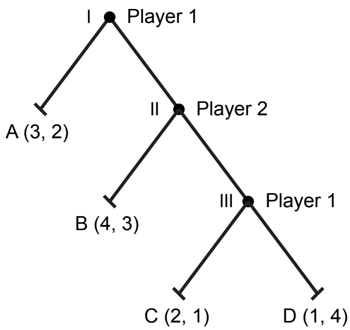

Figure 6.1: An extensive form representation of a two-player sequential game. Player 1 decides first, Player 2 second, and Player 1, again, third. The decision points are indicated in Roman numerals (I – III). Each end-node has a pair of payoffs, of which the left-side is Player 1's payoff and the right-side Player 2's payoff. Each player's goal is to obtain their highest possible payoff. In this particular game, the highest possible payoff for Player 1 is a 4, which is obtainable because Player 2's highest possible payoff is located at the same end node (i.e., B). Player 2's payoff of 4 is not obtainable because Player 1 would decide "left" instead of "right" at the third decision point (III).

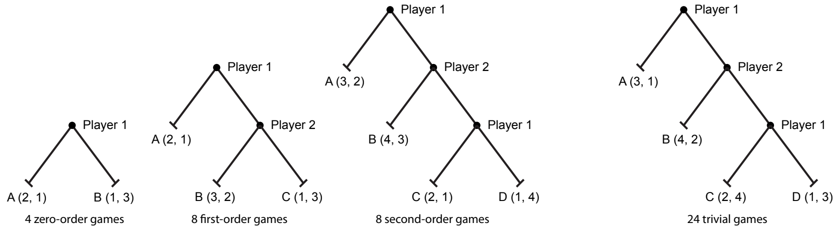

Figure 6.2: Extensive forms of example games (see Figure 6.1 for a detailed explanation). Stepwise Training consisted of 4 zero-order, 8 first-order, and 8 second-order games. Undifferentiated Training consisted of 24 trivial games. Each game had a unique distribution of payoffs.

training procedure, because it consisted of so-called trivial games (Figure 6.2; right panel), which are easier to play than truly second-order games such as in Figure 6.1. Trivial games are easier because Player 2 does not have to reason about Player 1's decision at III: Player 2's payoff in B is either lower or higher than both his payoffs in C and D. Consequently, Player 2 does not have to apply ToM, and Player 1 can suffice with first-order ToM. Flobbe et al. therefore argued that the training of Hedden and Zhang does not prepare people to play truly second-order ToM games. To test this claim, Meijering et al. administered two types of training procedures.

One group of participants was administered Hedden and Zhang's training procedure, which will henceforth be referred to as Undifferentiated Training, as all games had three decision points. The other group was administered Flobbe et al.'s training phase, but slightly modified (cf. Meijering et al., 2011). The latter training procedure will henceforth be referred to as Stepwise Training, as each additional decision point was introduced in subsequent blocks of games (Figure 6.2; left panel). Meijering et al. hypothesized that these training procedures would have distinct effects on strategy formation and thus performance. They predicted that Stepwise Training would facilitate participants to incorporate mental states of increasing complexity into their decision making process, yielding high accuracy. Undifferentiated Training, in contrast, would not motivate participants to develop recursive ToM, as they could suffice with application of first-order ToM. As expected, the participants that were assigned to Stepwise Training performed better than the participants assigned to Undifferentiated Training (see Figure 6.5).

One specific behavioral pattern is of particular interest to validate the model: The performance of participants assigned to Undifferentiated Training rose to ceiling during the training phase and dropped again when the experimental phase started (Figure 6.5). We hypothesize that the participants applied simple, child-like strategies during the training phase, because these strategies worked and did not consume much cognitive resources. At the start of the experimental phase, however, these strategies did not work anymore and accuracy dropped, because the games, while superficially similar, required more complex reasoning. Nevertheless, accuracy increased again over the course of the experimental phase, as the participants were able to revise their strategies. We will show that our computational cognitive model can simulate this process: The model's most important characteristic is that the complexity of its reasoning gradually increases by repeatedly attributing its own (evolving) strategy to the other player.

Computational cognitive model

The model[1] is implemented in the ACT-R cognitive architecture (Anderson, 2007; Anderson et al., 2004). ACT-R comprises a production system, which executes if-else rules, and contains declarative knowledge, which is presented as memory representations, or so-called chunks. In addition, ACT-R also includes modules that simulate specific cognitive functions, such as vision and attention, declarative memory, motor processing, et cetera. The results of these simulations appear as chunks in the modules' associated buffers, which the model continually checks (and manipulates) by means of its production system. ACT-R imposes natural cognitive constraints, as buffers can hold just one chunk at a time, and production rules can only fire successively, whenever their pre-specified conditions are matched. ACT-R does allow for parallel processing whenever a task induces cognitive processing in distinct modules. The model that we present here runs atop of ACT-R.

The model's behavior partially depends on memory dynamics. It needs to retrieve factual knowledge from declarative memory, and both the speed and success of retrieval depend on the so-called base-level activation of a fact (or chunk). The higher the base-level activation is, the greater the probability and speed of retrieval. The base-level activation in turn is positively correlated with the number of times a fact is retrieved from memory and the recency of the last retrieval.

The model simulates inference of mental states in sequential games. It uses a simple strategy at first and gradually revises that strategy until it can process recursive mental states. We consider the application of a particular strategy, and revising that strategy, to be deliberate processes. Therefore, application and revision are implemented by means of an interaction between factual knowledge and problem solving skills. Arslan, Taatgen, and Verbrugge (2013) successfully used a similar approach in modeling the development of second-order ToM in another ToM paradigm (i.e., the false-belief task). Van Rijn, Van Someren, and Van der Maas (2003) have successfully modeled children's developmental transitions on the balance scale task in a similar vein. Factual knowledge is represented by chunks in declarative memory, which store what strategy the model should be using. The problem solving skills, or strategy levels, are executed by (recursively) applying a small set of production rules. The model's goal is to make decisions that yield the greatest possible payoff. Decisions are either 'stop the game' or 'continue it to the next decision'. The model was presented with the same distributions of payoffs (i.e., items) as were presented to the participants.

The model's initial simple strategy is to consider only its own decision at the first decision point and to disregard any future decisions. The model's decision is based on a comparison between its (i.e., Player 1's) payoff in A and the maximum of its payoffs in B, C, and D. If the model's payoff in A is greater, the model will decide to stop. Otherwise, the model will decide to continue. By using this simple strategy the model seeks to maximize its own payoff, which can be considered a direct translation of the instructions given to the participants.

This strategy will work in some games but not in all. Whenever the strategy works, the model receives positive feedback and stores in declarative memory what strategy it is currently using. In fact, the model stores a strategy level, which is level-0 in the case of the

1 *The model can be downloaded from http://www.ai.rug.nl/~meijering/iccm2013*

simple strategy described above. Whenever the strategy does not work, the model receives negative feedback and stores in declarative memory that it should be using a higher strategy level (e.g., level-1).

The higher strategy level means that the model should attribute whatever strategy it was using previously to the other player at the next decision point. In the case of strategy level-1, the model attributes the model's initial simple strategy (i.e., level-0) to Player 2. Accordingly, the model is applying first-order ToM, as it reasons about the mental state of Player 2, who considers only his own payoffs and disregards any future decisions.

Again, this strategy will work in some games but not in all. Whenever it does not work, the model receives negative feedback and stores in declarative memory that it should be using a higher strategy level. At a higher strategy level, the model will attribute whatever strategy level it was using previously to Player 2. At strategy level-2, the model attributes strategy level-1 to Player 2, who in turn will attribute strategy level-0 to the player deciding at third decision point: Player 1. Now the model is applying second-order ToM.

Assumptions

The model is based on two assumptions. The first assumption is that participants, unfamiliar with sequential games, start playing according to a simple strategy that consists of one comparison only: Participants compare their current payoff, when stopping the game, against the maximum of all their future payoffs, when continuing the game. This strategy can be considered the simplest possible strategy, as participants who are using it ignore the consequences of any possible future decision, whether their own or the other player's.

If participants obtain expected outcomes, they do not have to revise their strategy. However, if participants obtain unexpected outcomes, they have to acknowledge that the unexpected turn of events was caused by the other player deciding at the next decision point. Reasoning about the other player, participants can only attribute a strategy they are familiar with themselves. This is our second assumption, which is based on variable frame theory (Bacharach & Stahl, 2000). Imagine a scenario in which two persons are asked to select the same object from a set of objects with differing shapes and colors but one person is completely colorblind. The colorblind person cannot distinguish the objects based on color, nor can he predict how the other would do that. Therefore, the colorblind person can only predict or guess what object the other would select based on which shape is the least abundant. The seeing person should account for the colorblind person's reasoning and also choose the object with the least abundant shape. This variable frame principle also applies to reasoning about others: We can only attribute to others goals, intentions, beliefs, and strategies that we are familiar with ourselves.

Mechanisms

The simple strategy is implemented in two production rules. The first production rule determines what the payoff will be when stopping the game; the other production rule determines what the highest future payoff could possibly be when continuing the game. Both productions are executed from the perspective of whichever player is currently deciding (Figure 6.3). The model will attribute this simple strategy from the current decision point to

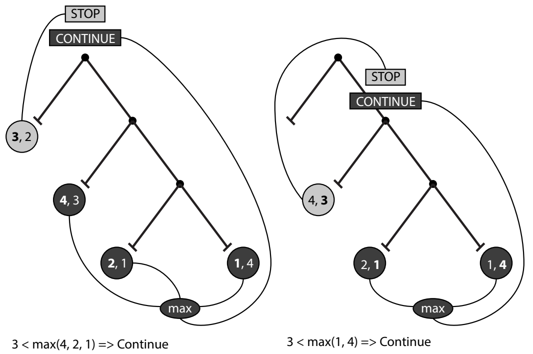

Figure 6.3: Depiction of the simple strategy. In the left panel, the model compares its payoff if it would stop (light grey) against its maximum possible payoff if it would continue (dark grey). In the right panel, the model compares Player 2's payoff if Player 2 would stop (light grey), against Player 2's maximum possible future payoff (dark grey). The left panel schematically represents the application of zero-order ToM, and the right panel the attribution of zero-order ToM to the other player.

the next, each time the model updates its strategy level (i.e., incrementing strategy level by one). The model will thus heighten its level, or order, of ToM reasoning.

Zero-order ToM

Before the model starts applying its strategy, it needs to construct a game state representation to store the payoffs that are associated with a *stop* and *continue* decision, respectively. To construct a game state, the model first retrieves from declarative memory what strategy level it is currently using. At the beginning of the experiment, strategy level has a value of 0, which represents the simple strategy. After retrieving strategy level, the model constructs its current game state.

Starting with the simple strategy, the model will determine its own *stop* and *continue* payoffs (see Figure 6.3, left panel), which will be stored in the game state representation. The model will then compare these payoffs and make a decision. After the model has made a decision, it will update declarative memory by storing what strategy level the model should be playing in the next game: If the model's decision was correct, the model should continue playing its current strategy level; otherwise the model should be playing a higher strategy level.

After playing a couple of games in which the simple strategy (i.e., level-0) does not work, the higher strategy level (i.e., level-1) will have a greater probability of being retrieved, as its base-level activation increases more than the simple strategy's base-level activation. At the start of the next few games, before the model constructs its game state, it will begin retrieving strategy level-1 from declarative memory.

First-order ToM

Playing strategy level-1, the model will first determine what payoff is associated with a *stop* decision at the first decision point (I). However, before determining what payoff is associated with a *continue* decision, the model needs to reason about the future and therefore consider the next decision point (II). It attributes strategy level-0 to Player 2, who is deciding at II.

Later, the model will return to the first decision point and determine what payoff is associated with a *continue* decision.

At II, the model will apply strategy level-0, but from the perspective of Player 2 (Figure 6.3, right panel). When reasoning about Player 2's decision, the model constructs a new game state, which references the previous one. The previous games state is referenced, because the model needs to jump back to that game state and determine what payoff is associated with a *continue* decision in that game state. At II, the model will execute the same production rules that it executed before when it was playing according to strategy level-0: It will determine what payoffs are associated with a *stop* and a *continue* decision, but from the perspective of Player 2.

The model will not produce a response whenever it determines the *stop* and *continue* payoffs at II, because the problem state at II references a previous one (i.e., I). The model will therefore backtrack to the previous game state representation, which did not yet have a payoff associated with a *continue* decision. That payoff can now be determined based on the current game state (i.e., Player 2's decision). The model will retrieve the previous game state from declarative memory.

After retrieving the previous game state representation, the model has two game states stored in two separate locations, or buffers: The current game state is stored in working memory, or the *problem state buffer* (Anderson, 2007; Borst, Taatgen, & Van Rijn, 2010), and the previous game state is stored in the *retrieval buffer*, which belongs to the declarative memory module. The model will determine what payoff is associated with a *continue* decision in the previous game state (stored in the retrieval buffer) given the decision based on the current game state (in the problem state buffer). It will update the previous game state and store it in working memory.

Playing strategy level-1 and being back in the previous game state, there is no reference to any previous game state and the model will make a decision based on a comparison between the payoffs associated with the *stop* and *continue* decisions. As explained previously, the model will stop if the payoff associated with stopping is greater; otherwise the model will continue.

Again, after the model has made a decision, it will update declarative memory by storing what strategy level the model should be playing in the next game(s). If the model's decision is correct, it will apply the current strategy level. Otherwise, the model will revise its strategy level by storing in declarative memory that it should be using strategy level-2 in the next game(s).

Second-order ToM

The model will first determine what payoff is associated with stopping the game and then consider the next decision point. There, the model proceeds as if it were playing strategy level-1, but from the perspective of Player 2. In other words, the model is applying second-order ToM.

The strategy described above closely fits the strategy of forward reasoning plus backtracking (Meijering, Van Rijn, Taatgen, & Verbrugge, 2012; Chapter 5 in this dissertation). Meijering et al. (2012) conducted an eye-tracking study, and participants' eye movements reflected a forward progression of comparisons between payoffs, followed by backtracking to previous decision points and payoffs. Such forward and backward successions are present in strategy level-2 as well: Payoffs of *stop* decisions are determined one decision point after another, and this forward succession of payoff valuations is followed by backtracking, as payoffs of previous *continue* decisions are determined in backward succession.

Results

The model was presented with the same trials as in Meijering et al.'s (2011) study (see also Chapter 2 in this dissertation), with stepwise training versus undifferentiated training as a between-subjects factor. The model was run 100 times for each training condition. Each model run consisted of 20 (stepwise) or 24 (undifferentiated) training games, followed by 64 truly second-order games. The results are presented in Figures 6.4 and 6.5.

Figure 6.4 shows the proportions of models that apply strategy levels 0, 1, and 2, calculated per trial. The left panel of Figure 6.4 shows the output of the models that received 24 undifferentiated training games before playing 64 second-order games. As can be seen, initially all models apply strategy level-0, corresponding with zero-order ToM, but that proportion decreases quickly in the first couple of games. The proportion of models applying zero-order ToM decreases because that strategy yields too many errors, which can be seen in Figure 6.5. The models store in declarative memory that they should be using strategy level-1, but it takes a few games before the base-level activation of the level-0 chunk drops below the retrieval threshold. After it does, the models start retrieving level-1 chunks and will apply strategy level-1, which corresponds with first-order ToM. The proportion of models that use strategy level-1 increases up to 100% towards the end of the 24 undifferentiated training games. The models do not start applying strategy level-2 during the training phase, because strategy level-1 yields correct decisions in all undifferentiated training games, which can be seen in Figure 6.5. However, in the experimental games, which are truly second-order games, strategy level-1 yields too many errors, and accuracy drops. It takes approximately 40 games before the base-level activation of the level-1 chunk has dropped below the threshold in at least half of the models. The models gradually start using strategy level-2, and accuracy starts to increase again, as can be seen in Figure 6.5.

The right panel of Figure 6.4 shows the output of the models that were presented with 20 stepwise training games (4 zero-order, 8 first-order, and 8 second-order games) before playing

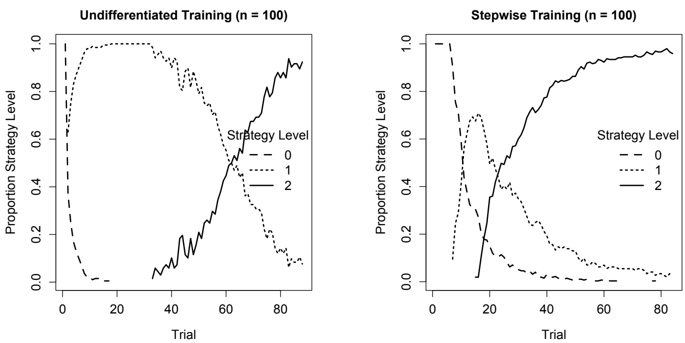

Figure 6.4: Proportion of models that apply strategy levels 0, 1, and 2; plotted as a function of trial. The left panel depicts these proportions for the model that received undifferentiated training; the right panel depicts the proportions for the model that received stepwise training.

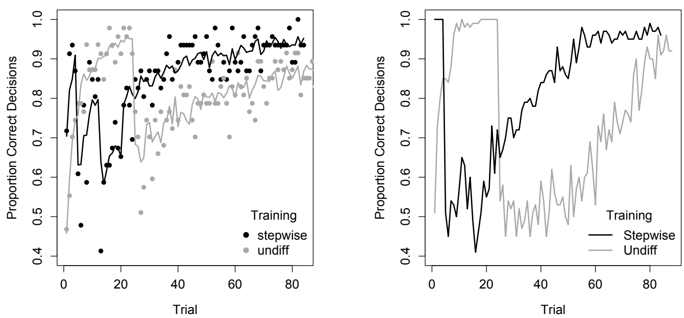

Figure 6.5: Proportion of correct decisions, or accuracy, across participants (left panel) and models (right panel). The solid lines in the left panel represent the fit of the statistical model, which is added to visualize the proportion trends.

64 second-order games during the experimental phase. As can be seen, all models start applying strategy level-0, and they use it longer than the models that received undifferentiated training. The reason is that strategy level-0 yields a correct answer in the first four games during stepwise training, because those are zero-order games. As can be seen in Figure 6.5 (right panel), accuracy is 100% in the first few games. In the next eight first-order training games (Trials 5 – 12), the proportion of models that apply strategy level-0 decreases, as strategy level-0 yields too many errors. Simultaneously, the proportion of models applying strategy level-1 increases, as the base-level activation of the level-0 chunk decreases and the models start retrieving the level-1 chunk. In the next eight second-order training games (Trials 13 – 20), the proportions of models that apply strategy level-0 and level-1 decrease, as both strategy levels yield too many errors. Simultaneously, the proportion of models that apply strategy level-2 increases. As strategy level-2 yields a correct decision in the remainder of the games, accuracy increases up to ceiling, which can be seen in Figure 6.5 (right panel).

The accuracy trends in the models' output qualitatively fit those of Meijering et al.'s study (2011). The quantitative differences are probably due to the fact that not all participants started out using the simple strategy, whereas all models did. One possible explanation is that some participants started with intermediate-level strategies and, due to large proportions of optimal outcomes, did not proceed to the highest level of reasoning. We could account for this by storing level-0, level-1, and level-2 chunks in declarative memory, and having the base-level activation of these chunks follow the distribution of zero-order, first-order, and second-order ToM in the adult population. A meta-review of (higher-order) ToM in adults and children may be a good starting point to find the appropriate distributions. Nevertheless, the qualitative trends in the model data, changing as a function of game complexity, correspond with the response patterns in the behavioral data. The trends suggest that people use simple strategies for as long as these yield expected outcomes.

In the introduction we hypothesized that simple strategies are a legacy of our childhood years, and that adults keep using those strategies that have proven themselves successful during development. To test this hypothesis, we have re-analyzed the data from Flobbe et al.'s (2008) developmental study. We expected that few children would have sufficient cognitive resources to apply second-order ToM, and that performance levels would therefore align well with lower and intermediate strategy levels. The most obvious prediction is that prevalence of level-0, level-1, and level-2 strategies can be ranked, where level-0 is the most dominant strategy and level-2 is least frequent.

Developmental study

Flobbe et al. (2008) studied the application of second-order ToM in children that were in between 8 and 10 years (M = 9;2). They presented the children with sequential games, and performance was just above chance-level (57% correct). As children of age 9 are at the brink of mastering second-order ToM (Flobbe et al., 2008; Miller, 2009; Perner & Wimmer, 1985), we expect the lower and intermediate strategies to be most prevalent in Flobbe et al.'s study, which is thus perfect to validate our model.

We hypothesize that children apply the same simple strategies that are implemented in our computational cognitive model. We predict that the children start out with the simplest (i.e., zero-order) strategy, and that some will learn to attribute that strategy to the other player. Probably few children will learn that the other player, in turn, attributes the simple strategy to the player who decides next (i.e., to them). As each child was first asked to predict the other player's decision, before they were asked to make a decision themselves, we have a direct measure of the child's perspective of the other player's strategy. We will analyze both the predictions and the decisions.

Predictions

We applied a binomial criterion to reliably categorize a participant's predictions as belonging to either level-1 or level-2: The predictions in at least 8 out of 10 consecutive games had to be congruent with one particular strategy level to label the predictions accordingly. This might seem strict, but 8 out of 10 is the minimum quantile that is still significant with a significance level of 0.05. As the experiment consisted of 40 second-order games, we categorized each child's responses in 4 sets of 10 games. Figure 6.6 depicts the proportion of children that applied either first-order or second-order ToM. These ToM-orders correspond with level-1 and level-2 in the computational model.

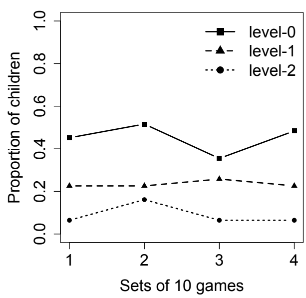

Figure 6.6: The proportion of children that applied zero-order ToM (level-0), first-order ToM (level-1), or second-order ToM (level-2) to the other player; depicted in 4 consecutive sets of 10 games.

Note that sets of predictions that could not be categorized level-1 or level-2 do not necessarily imply the use of level-0, because the predictions in those sets could have been completely random, or a mixture of the various strategy levels. The decisions are therefore analyzed to determine the prevalence of strategy level-0.

As can be seen in Figure 6.6, the proportion of children that applied first-order ToM by attributing strategy level-0 to the other player is greater than the proportion of children that applied strategy second-order ToM. Furthermore, many children's predictions could not be labeled according to one of the strategies at all (13 out of 40). These children probably switched frequently between multiple possible perspectives, and such switching is difficult to reliably capture by means of a statistical model. Nevertheless, most of the children whose responses could be categorized, were applying first-order ToM by attributing the simple (i.e., level-0) strategy to the other player. Almost none of the children was able to consistently attribute strategy level-1 to the other player, thereby applying second-order ToM.

Decisions

As explained above, the predictions required application of first-order ToM at minimum and could therefore not be indicative of zero-order ToM. Therefore, the decisions were analyzed to determine how many children applied zero-order ToM, ignoring the other player entirely. Again, we categorized the decisions based on the binomial criterion that at least 8 out of 10 consecutive responses should be consistent with application of zero-order ToM (i.e., level-0 in the model). As can be seen in Figure 6.6, most of the children that consistently responded according to one of the strategies applied zero-order ToM when making a decision. This is remarkable, because each child that participated in the experimental phase successfully passed a training block in which they were required to apply first-order ToM. This finding suggests that the children could not see how first-order ToM would fit in the more complex games in the experimental blocks. They may have recognized that it did not work, but still could not revise their strategy to incorporate an additional ToM level.

To conclude, a re-analysis of Flobbe et al.'s (2008) study shows that few children were able to apply second-order ToM (level-2), and that most children used simple strategies. The most dominant strategy was the simplest one that did not account for any future decision points. Most children seemed to apply zero-order ToM (level-0) while making a decision. Some children, though, were able to attribute that simple strategy to the other player, thereby applying first-order ToM (level-1). These strategies are the same as those implemented in our computational cognitive model. The model is thus supported in two ways: (1) Its most simple strategies are found in children, and (2) it learns to revise its strategies as adults do.

Conclusions

In this study we presented a computational cognitive model that simulates inference of mental states in sequential games. More specifically, the model was required to apply ToM recursively, a skill that appears to be unique to human intelligence. Many studies have shown that people oftentimes fail to apply ToM to interpret the behavior of others (e.g., Apperly et al., 2010; Keysar et al., 2003; Lin et al., 2010). In this study, in contrast, we show that people do not necessarily fail to apply ToM, but rather first apply simple strategies that are computationally less costly. Only when necessary do people revise their strategies to account for complex mental states.

The model is based on previous empirical findings (Meijering et al., 2011) that seemed to imply that people exploit the possibility of using simple strategies for as long as these pay off. We implemented one such simple strategy that ignores any future decisions and simply compares the immediate payoff, when stopping a game, against the maximum of all future possible payoffs. By means of simple memory dynamics the model either retrieves a chunk that specifies that the model should continue using this strategy, or chunks that specify that the model should attribute the simple strategy to the player who decides next. Although this updating process may seem simplistic at first sight, the model does gradually master second-order ToM, but only because that is required in the games in this study. In other words, the model's most important dynamics are not task-specific, and because of that, the model is flexible and can accommodate many other two-player sequential games.

We found support for the model in the data from Flobbe et al.'s (2008) developmental study in which 9-year-old children were presented with similar sequential games. Most children used the simple, level-0, strategy when making a decision. The second-most prevalent strategy was the level-1 strategy. Using that strategy, the children attributed the simplest possible strategy (i.e., level-0) to the other player. Few children were able to apply second-order ToM mind. They did not recognize that the other player, in turn, attributed the simplest strategy (i.e., level-0) to them. These findings show that the children used the same simple strategies as the adults initially used in Meijering et al.'s study. However, the adults were able to revise their strategies to achieve the highest required level of ToM reasoning, whereas the children may not have had sufficient cognitive resources to achieve that same level of reasoning.

Our notion of zero-order ToM (i.e., strategy level-0) closely maps with the instruction given to the participants: to maximize their payoff. This strategy corresponds with a risk-seeking perspective, because it does not account for the fact whether higher future payoffs are actually attainable. There are other notions of a level-0 strategy, however. A risk-seeking strategy can be contrasted with a risk-aversive strategy according to which one would stop if there were any lower future payoffs. There is still another notion of a level-0 strategy: Hedden and Zhang (2002; 2012) defined a so-called myopic level-0 strategy that only considers the current payoff and the closest future payoff. Player 1, for example, would only compare his payoffs in A and B, ignoring his payoffs in C and D. These strategies, however, are almost non-existent in Flobbe et al.'s dataset.

The findings from this study raise the question why younger children of 6 to 8 years are perfectly capable of accounting for second-order mental states in traditional false-belief studies (Coull, Leekam, & Bennett, 2006; Flobbe et al., 2008; Perner & Wimmer, 1985; Sullivan, Zaitchik, & Tager-Flusberg, 1994), as well as when they are asked to discriminate between ironic and deceptive speech acts (Winner & Leekam, 1991). One possible explanation is practice: Children have encountered false beliefs, irony, and deception more often than games such as in this study. Another explanation is that games can have a large space of possible outcomes, which requires extensive reasoning. False-belief stories and speech acts, on the other hand, are a given and thus require fewer computations. On a related note, children are better at reasoning about past events than about future possible outcomes (e.g., McColgan & McCormack, 2008; Suddendorf, Nielsen, & Gehlen, 2011). Reasoning about past events can be considered a linear traversal backwards in time, whereas reasoning about future events may follow an expanding tree-like structure.

This study has at least two methodological implications: One, experimenters should be careful in selecting 'practice' items, as participants exploit the possibility of using simple strategies when possible. Two, average proportions of correct answers, a popular statistic in most ToM studies, may not be as informative as a categorization of responses (also see Raijmakers et al., 2013). Flobbe et al., for example, reported that performance was just above chance-level (i.e., 57% correct), and the most common interpretation would be "on average children were able to apply second-order ToM in 57% of the games." However, the current study shows that this score can be obtained if 1 or 2 children are applying second-order ToM and most of them below-optimal strategies such as zero-order and first-order ToM.

The theoretical implication of this study is that people do not necessarily perceive sequential games in terms of interactions between mental states. They know that there is another player making decisions, but they have to learn over time, by playing many games, that the other player's depth of reasoning could be greater than initially thought. Learning takes place when people obtain unexpected outcomes and start recognizing that the other player has a role in their outcomes. They will have to attribute their own, simple, strategies to the other player, thereby developing increasingly more complex strategies themselves. Over time, reasoning will become as complex as necessary, as simple as possible.