Are you thinking what I'm thinking? An Evaluation of a Simplified Theory of Mind

David V. Pynadath[1], Ning Wang[2], and Stacy C. Marsella[1]

[1] University of Southern California, Institute for Creative Technologies
12015 E Waterfront Dr, Los Angeles, CA 90094 USA
[2] Curious Lab LLC
7430 Arizona Ave, Suite 35, Los Angeles, CA 90045 USA

Abstract. We examine the effectiveness of an agent's approximate theory of mind when interacting with human players in a wartime negotiation game. We first measure how accurately the agent's theory of mind captured the players' actual behavior. We observe significant overlap between the players' behavior and the agents' idealized expectations, but we also observe significant deviations. Forming an incorrect expectation about a person is not inherently damaging, so we then analyzed how different deviations affected the game outcomes. We observe that many classes of inaccuracy in the agent's theory of mind did not hurt the agent's performance and, in fact, some of them played to the agent's benefit. The results suggest potential advantages to giving an agent a computational model of theory of mind that is overly simplified, especially as a first step when investigating a domain with as much uncertainty as wartime negotiation.

Keywords: theory of mind, cognitive models, wartime negotiation, evaluation of formal models

1 Introduction

Theory of mind is critical for success in social interaction [20]. Without it, people would not be able to understand each other's perspectives and desires. With it, people can form expectations of the others' behavior and choose their own behaviors informed by those expectations. Consequently, intelligent virtual agents (IVAs) also need theory of mind to successfully interact with people [5, 9].

Researchers have used a variety of cognitive models to provide agents with this capability. A common approach for realizing theory of mind is for the agent to use its own reasoning mechanism as a model for the reasoning of others, after substituting the others' beliefs, goals, capabilities, etc. for its own. For example, an agent that uses partially observable Markov decision problems (POMDPs) [6] for its own decision-making can model other agents and people as using POMDPs of their own [1, 4, 10, 14].

One of the challenges in implementing theory of mind within an IVA is its unavoidable uncertainty about the mental states of the others. For instance, even

Proceedings of the International Conference on Intelligent Virtual Agents, 2013.
© Springer-Verlag Berlin Heidelberg 2013

when playing a simple game like the Prisoner's Dilemma, an agent can never be sure what goals a human player is pursuing. An agent with a completely accurate theory of mind would have to capture whether players care about maximizing only their own payoff, or social welfare, fairness, and many other payoffs as well, not to mention the possible expectations they have about the agent's own behavior. In reality, a agent rarely has an accurate method for assessing the probability of all of these alternate hypotheses, so it is likely to make errors when making decisions based on these inaccurate assessments. Furthermore, there are computational costs associated with maintaining a richer, more accurate model, costs that can be hard to justify if the accuracy does not benefit the agent's own utility [15]. An IVA has to carefully balance this trade-off between complexity and accuracy when deciding how rich a theory of mind to use.

In this paper, we examine the effectiveness of an agent's theory of mind when interacting with people in a wartime negotiation game [16]. The game pits human players against agents implemented in PsychSim, a computational architecture for theory of mind that has been used to build IVAs in other negotiation domains [8, 9, 11]. The models use asymmetry of information to provide only one side (e.g., an agent) with complete information about the game's underlying likelihoods and costs, while leaving the other side (e.g., a human player) in uncertainty about those parameters. While the agent has the advantage of complete information, its performance is highly dependent on its ability to perform effective theory-of-mind reasoning about the human player's uncertainty. In our study, the agent's simplifying assumption that the player had no such uncertainty helped to achieve the initial goal of evaluating the coarse game-theoretic predictions of human behavior [16]. In this paper, we further analyze the data to study the impact the agent's approximate theory of mind had on its performance.

We begin our investigation of the agent's complexity/accuracy trade-off by measuring the accuracy of the agent's theory of mind. We observe significant overlap between the human players' actual behavior and the agents' idealized expectations, but we also observe significant deviations. Forming an incorrect expectation about a person is not *inherently* damaging, so we then analyzed how different deviations affected the game outcomes. We observe that many classes of inaccuracy in the agent's theory of mind did not hurt the agent's performance and, in fact, some of them played to the agent's *benefit*. The results suggest potential advantages to giving an agent a computational model of theory of mind that is overly simplified, especially as a first step when investigating a domain with as much uncertainty as wartime negotiation.

2 Wartime Negotiation

A number of formal models in the political science literature represent war as a costly process embedded within a negotiation game. In these models, two sides are in a dispute over the division of a desirable resource, which we will illustrate as territory claimed by both sides. The game begins with some initial split of the territory. The game progresses round by round, with each round consisting of one

side proposing a split of the territory, the other side responding to that proposal, and a possible battle between the two. The game ends with a final split achieved by either an agreement on the proposed split or a decisive military victory by one side on the battlefield.

We chose two models, *Powell* [13] and *Slantchev* [18], for this investigation, based on their impact on the field and their appropriateness for a human-agent game interaction. Both models assume fixed probabilities associated with the battlefield, so that one side's probability of winning does not change during the course of the game, regardless of previous military outcomes. The costs of a single battle are also fixed throughout the course of the game. In our study, we present these costs to the human players in terms of troops lost.

A critical property of these models is uncertainty about the costs to the other side and the likelihood of battlefield outcomes. If both sides had complete information about the costs and probabilities, they could do an exact cost-benefit analysis and immediately agree upon a territorial split. In both models we implemented, only one side has complete information and the other is uncertain about the probability and costs of battlefield outcomes. This asymmetry lends itself to our human participant study, as we can give the agent complete information about the game probabilities and costs, but withhold that information from the human player. Even with complete information about the *game* uncertainty, the agent still needs to model the *players'* uncertainty and how it affects their decisions.

2.1 The Powell Model

In the following *Powell* model [13], Player 1 is a human player and Player 2 is an agent:

1. Player 1 makes an offer of $x\%$ of the territory.
2. Player 2 decides to accept, reject, or attack.
   (a) If accept, Player 2 gets $x\%$, Player 1 gets $(100 - x)\%$, and game ends.
   (b) If attack, Players 1 and 2 lose c_1 and c_2 troops, respectively. Player 1 collapses with probability p_1 and Player 2 collapses with probability p_2.
      i. If only Player 1 collapses, Player 2 gets 100%, and game ends.
      ii. If only Player 2 collapses, Player 1 gets 100%, and game ends.
      iii. Otherwise, return to Step 1.
   (c) If reject, Player 1 decides whether or not to attack.
      i. If attack, go to Step 2b; otherwise, return to Step 1.

Player 1 (the human player) does not have prior knowledge of the probabilities of collapse (p_i) or the costs of war (c_i), but Player 2 (the agent) does. In the game-theoretic analysis of this model, Player 2 can compute the optimal threshold, where it should accept any offer from Player 1 that exceeds the threshold and reject or attack otherwise. This threshold is lower the higher its costs (c_2), the higher its probability of collapse (p_2), and the lower Player 1's probability of collapse (p_1). Because Player 1 does not know these values, it is uncertain about

Player 2's threshold for accepting an offer. The equilibrium behavior can be described as *screening*, where Player 1 will make a series of increasingly attractive offers, expecting weaker opponents (who have a lower threshold) to accept early in the process, thus screening them out before making the higher offers necessary to appease stronger opponents [13].

2.2 The Slantchev Model

Unlike the Powell model's probability of collapse, the *Slantchev* model includes an additional variable, *military position* ($k \in \{0, 1, 2, \ldots, N\}$), that represents gradual military progress toward or away from complete collapse [18]. We again have a human as Player 1 and an agent as Player 2:

1. The initiating player makes an offer of $x\%$ of the territory.
2. The responding player decides to accept or reject the offer.
   (a) If accept, the responding player gets $x\%$, the initiating player gets $(100-x)\%$, and game ends.
   (b) If reject, continue to Step 3.
3. Battle occurs, and Players 1 and 2 lose c_1 and c_2 troops, respectively. Player 1 wins the battle with probability p, Player 2 with probability $1-p$.
   (a) If Player 1 wins, $k \leftarrow k+1$. If $k = N$, then Player 1 gets 100% and game ends.
   (b) If Player 2 wins, $k \leftarrow k-1$. If $k = 0$, then Player 2 gets 100% and game ends.
4. Return to Step 1 with initiating and responding players reversed.

Like the Powell model, Player 1 does not know the battle probability (p) or costs (c_i), but Player 2 does. Thus, the agent can compute a threshold for acceptable offers, but this threshold is now a function of k, the current military position. This threshold increases (and the agent's counteroffers decrease) as k decreases, c_2 decreases, and p decreases. As in the Powell model, the human players are uncertain about this threshold because of their ignorance of the underlying probability and cost, so the equilibrium behavior again exhibits some screening.

Examining human behavior in these two games allows us to study the effectiveness of the agent's theory of mind. In addition to the variation between our two game models, we also vary the players' starting territory (28% vs. 72%) to possibly shift their reference points in the negotiation [7, 12]. These four combinations produce the four experimental conditions summarized in Table 1.

3 PsychSim Agents in Wartime Negotiation

We implemented both the Powell and Slantchev games within PsychSim, a multiagent framework for social simulation [10, 14]. PsychSim agents have their own goals, private beliefs, and mental models about other agents. They generate

	Powell28	**Powell72**	**Slantchev28**	**Slantchev72**
Start	28%	72%	28%	72%
War	A single battle can end the war in a win/loss for the player		A single battle changes military position $\in [0, 10]$, with 10 (0) being a win (loss) for the player	
Battle	Battle occurs only if either side unilaterally initiates		Battle occurs every round, *not* initiated by either side	
Offers	Agent cannot counteroffer		Agent must counteroffer	

Table 1: Game features across the four experimental conditions

their beliefs and behaviors by solving POMDPs [6], whose quantitative transition probabilities and reward functions capture the game-theoretic dynamics of our chosen models of wartime negotiation as follows:

State: Territory (0–100%), number of troops, military position (Slantchev only)
Actions: Accept/reject offer, attack (Powell), offer player x% of territory (Slantchev)
Transition: The probability distribution of the effects of actions on the state
Observation: We assume that the agent has complete information
Reward: Linear in amount of territory and number of troops

The PsychSim agent's theory of mind expands this POMDP model to include the human players' POMDP models as well. The agent does not know what POMDP would make the best model of the human player. For example, the game does not reveal the probability of battlefield outcomes to the player's side, so it is not clear what transition function would best capture the player's expectations. Furthermore, while the player most likely wants to increase both territory and surviving troops, the agent has no way of knowing whether that desire fits a linear function like its own, let alone what weights to use even if it does. On top of this uncertainty about the player's model of the game, the agent's theory of mind must also capture the players' theory of mind about itself.

Rather than trying to capture all of this uncertainty within the agent's theory of mind (e.g., by giving it a distribution over multiple POMDP models of the player), we instead implemented an agent that has no such uncertainty. In particular, it models the human player as following a POMDP that has complete information just as its own does. This POMDP uses a linear reward function, just like the agent's, but increasing in the human player's territory and troops instead. Thus, the agent uses a fixed model of the player as following the optimal policy computed by solving this POMDP. The agents can then use that policy as an expectation within its own POMDP to compute an optimal policy for itself.

3.1 Powell agent behavior: Attacking

Before an agent can decide whether it is optimal to accept an offer or not under the Powell condition, it must first examine its subsequent choice of simply rejecting an offer, or else rejecting the offer and attacking the player[3]. If either side

[3] Under Slantchev, there is no such choice, as a battle occurs every round.

attacks, the agent will win the war with probability $p_1(1-p_2)$ and will lose the war with probability $p_2(1-p_1)$. For our study, we use $p_1 = p_2 = 10\%$, so both probabilities come out to 9%. If the agent attacks, it therefore expects to have 100% of the territory with 9% probability, 0% with 9% probability, and its original amount of territory with 82% probability. Regardless of the outcome, the agent will incur a fixed loss of troops that, in its reward function, is weighted the same as 2% of territory. Thus, the difference between an attack (by either side) and no attack is the 18% chance of a military resolution (with an expected split of either 0% or 100% for the agent) and the 2% cost in troops. This possibility is valued differently by the agent depending on its starting territory:

Powell28: When not accepting an offer, the agent does not attack
Powell72: When not accepting an offer, the agent always attacks.

In other words, the possible military resolution is appealing in the Powell72 game when the agent starts with only 28% of the territory. Furthermore, the agent's Powell72 policy implies that the players *never* get a chance to attack, as they get that choice only if the agent rejects without attacking.

3.2 Powell agent behavior: Accepting Offers

Because the agent will not attack in the Powell28 condition, it weighs the human player's offer against its current 72% of the territory. Under the Powell72 condition, on the other hand, the agent weighs the human player's offer against both its starting 28% territory *and* the possibility of earning a more favorable split when it attacks. Because the agent does not distinguish between earning territory at the negotiation table vs. on the battlefield, it is willing to hold out for a higher offer than just the status quo in Powell72. In particular, the POMDP solution generates the following thresholds for the agent:

Powell28: Accepts offers $\geq 71\%$ (rejects otherwise)
Powell72: Accepts offers $\geq 35\%$ (attacks otherwise)

3.3 Slantchev agent behavior: Accepting Offers

The agent performs a similar computation under Slantchev, except that battlefield expectations are now contingent on military position, $k \in \{0, \ldots, 10\}$. We set the probability, p, of a player winning a battle to be 30%, allowing the agent to compute its chances of winning or losing the war, which happens when $k = 0$ or 10, respectively. Solving the POMDP gives the agent a policy of holding out for higher offers as it gets closer to winning (i.e., k is low):

Slantchev28

$$\text{Accepts offers} \begin{cases} \geq 92\% \text{ if } k = 1 \\ \geq 85\% \text{ if } k = 2 \\ \geq 72\% \text{ if } k \in [3, 7] \\ \geq 66\% \text{ if } k = 8 \\ \geq 51\% \text{ if } k = 9 \end{cases}$$

Slantchev72

$$\text{Accepts offers} \begin{cases} \geq 79\% \text{ if } k = 1 \\ \geq 64\% \text{ if } k = 2 \\ \geq 30\% \text{ if } k \in [3, 7] \\ \geq 27\% \text{ if } k = 8 \\ \geq 20\% \text{ if } k = 9 \end{cases}$$

3.4 Slantchev agent behavior: Making Offers

To decide what offers it should make in the Slantchev game (it makes no offers under Powell), the agent first repeats the computation of Section 3.3 from the player's perspective. Again, it assumes that the players know that $p = 30\%$, and will have lower thresholds than the agent does because of this military disadvantage. It also does not change its beliefs about the player's thresholds, leading it to adopt a fixed policy of making offers as follows:

Slantchev28: Agent offers 10%, unless close to losing ($k > 7$), then offers 30%
Slantchev72: Agent offers 70%, unless close to winning, then offers 20% if $k = 2$, or 10% if $k = 1$.

4 Method

We recruited 240 participants, of an average age of 35, via Amazon Mechanical Turk. 51% of the participants are female and 49% are male. 65% of the participants are from the United States, 29% from India and 6% from other countries. 12% of the participants have some high school or high school diploma; 63% have some college or college degree and 25% have some graduate school or graduate degree. 13% of the participants use a computer for 1-4 hours a day, 43% use one 5-8 hours a day and 44% use one more than 8 hours a day.

Each participant is first assigned an anonymous ID and then reads the information sheet about the study. Then the participant fills out a Background Survey. Next the participant plays the negotiation game four times, each time with a different agent from one of the four conditions (the order is randomized). The game interface presents the participants with the troops and territory they own, as well as the number of rounds left and the history of previous offers and battle outcomes. There is no implication in the instructions that the participant would be playing against another human player. During the negotiation, the participant fills out an In-Game Survey. Following each negotiation game, the participant fills out an Opinion Survey. The study is designed to be completed within an hour, although the average duration was 32 minutes in our data.

We measured outcomes based on the following survey and game results:

Background Survey asks questions about the participant's age, gender, nationality, education, computer experience, Attitude Towards War [3], Social Orientation [19] and attitude towards Inappropriate Negotiation (SINS, from [17]).

Opinion Survey contains questions regarding the participant's goals during the game and modified questions from the Subjective Value Index (SVI) survey that measures perceptions of the negotiation outcome, process, relationship and the negotiator themselves [2].

In-Game Survey asks the participant to estimate the opponent's response after he/she makes an offer, e.g. accept the offer, reject it or attack.

Game Logs capture the actions the participant takes, the agent's actions and the world states, e.g. amount of troops and territory each side has.

5 Accuracy of the Agents' Theory of Mind

From our study's 240 participants, we have 238 games in the Powell72 condition and 239 games each in the Powell28, Slantchev72, and Slantchev28 conditions. We analyzed the participants' decisions made in all the games they played and categorized these decisions based on whether the agent would have made the same decision if it were in the participants' position. From Table 2, we see that the degree of conformity varies widely across the different conditions and decision types (empty cells are conditions where the action was inapplicable).

Player Action	Powell 28	Powell 72	Slantchev 28	Slantchev 72
Making Offers	84%	59%	21%	34%
Response to Offer	—	—	80.3%	75.8%
— Reject Offer	—	—	100%	100%
— Accept Offer	—	—	0%	0%
Decision to Attack	43%	—	—	—
— Attack	100%	—	—	—
— Not Attack	0%	—	—	—

Table 2: Percentage of participant decisions that matched agent's theory of mind

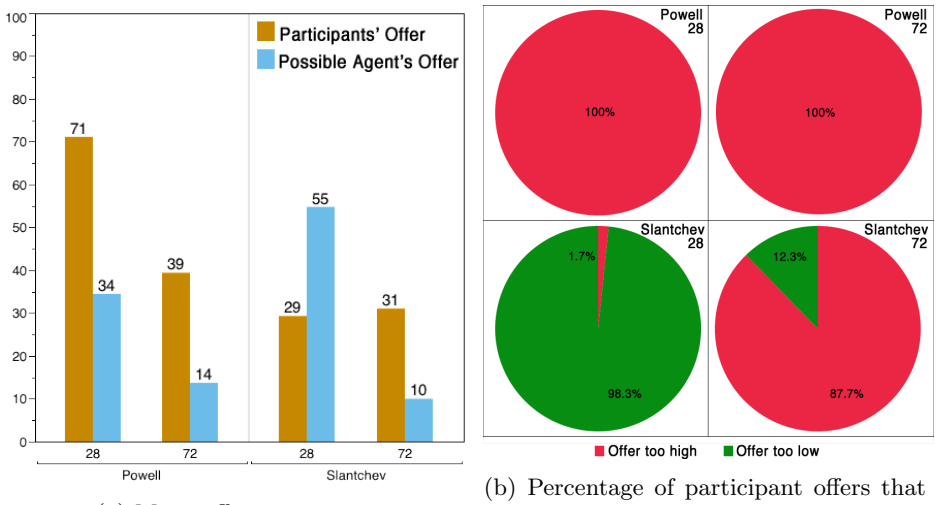
5.1 Participant behavior: Attacking

From Section 3.1, we know that the agent will attack instead of reject in Powell72, preempting the participants' potential choice to do so themselves. In Powell28, the agent makes the opposite choice, and its theory of mind leads it to expect the participants to act as it did in the same situation (i.e., having only 28% of the territory) and attack instead of reject. The participants' matched that expectation and decided to attack only 43% of the time. One possible explanation for this deviation is that the participants may have placed a higher value on troops than the agent's reward function did. Alternatively, because the participants do not know the probabilities of collapse, their uncertainty may have led them to underestimate their expected gains from battle.

5.2 Participant behavior: Making Offers

From Table 2, we see that the offers participants made did not always fall within the range of offers the agent would have made. Section 3.4 lays out the ranges of the agent's offers under Slantchev. In Powell games, the agent does not make offers, so we instead use the agent's offer-accepting policy (Section 3.2) as the range of offers the agent would have made. In particular, with a starting territory of 28% (for the agent), the agent will not offer more than 65%; with a starting territory of 72%, it will not offer more than 28%. One-way ANOVA tests show that, overall, the participants made offers consistent with the agent's policy more often under Powell than under Slantchev ($Mean_P = 77\%$, $Mean_S = 26\%$,

$p < .0001$). Within Powell games, as shown in Table 2, the participants' offers conformed more when starting with less territory ($p < .0001$). Within Slantchev games, however, the participants' offers conformed more when starting with more territory ($p < .0001$).



(a) Mean offer amounts.

(b) Percentage of participant offers that were higher/lower than the agent's.

Fig. 1: Offer amounts when participants deviated from the agent's policy.

To see how the participants' offers quantitatively differed from the agent's model, we calculated the offers the agent would have made given the same game states the participants were in. As a crude approximation, we calculated the agent's offer as a uniform distribution within the offer range. In Powell, for example, we modeled an agent starting with 28% of territory as calculating its offers from a uniform distribution from 0 to 65%. Figure 1a shows the participants' and agents' offers when a deviation occurred. A total of 756 offers made by the participants were included in the analysis, including 227 in Powell28 condition, 157 in Powell72, 273 in Slantchev28 and 99 in Slantchev72. Paired-sample t-tests show significant differences between the participants' and the agent's offers ($p < .0001$) in all four conditions. In the Powell games, participants made higher offers than the agent would have. Section 5.1 gave multiple hypotheses for why the participants might be less likely to attack than the Powell agent, and the same causes would also lead them to be overly generous to minimize the risk of war. Alternatively, the participants may have a preference for achieving a resolution through negotiation than through war, so they may bias their offers to increase the chance of acceptance by the Powell agent. In the Slantchev games, the participants' deviating offers exceed that of the agent under Slantchev72, but it is *lower* under Slantchev28. The contrast of Slantchev28 is even starker in Figure 1b, where we see the vast majority of deviating offers are lower than

the agent's. We do not have a good explanation for this outlying condition, although given the broad uncertainty facing the participants' they may be simply overestimating their probability of winning a war in this case.

5.3 Participant Behavior: Accepting Offers

The participants never receive offers to accept/reject under the Powell conditions, so we focus on only Slantchev games here. Table 2 shows that, when responding to an offer, the participants' responses match the agent's responses 80.3% of the time in Slantchev28 and 75.8% in Slantchev72. All of these deviations occur when the participants accepted an offer that the agent would not have. In other words, the agent would also have rejected any offer that the participants chose to do. This also implies that the agent never made offers that it would have accepted itself. Its theory of mind leads it to expect the participants to be more lenient given their military disadvantage, even though, in reality, the participants are initially unaware of that disadvantage.

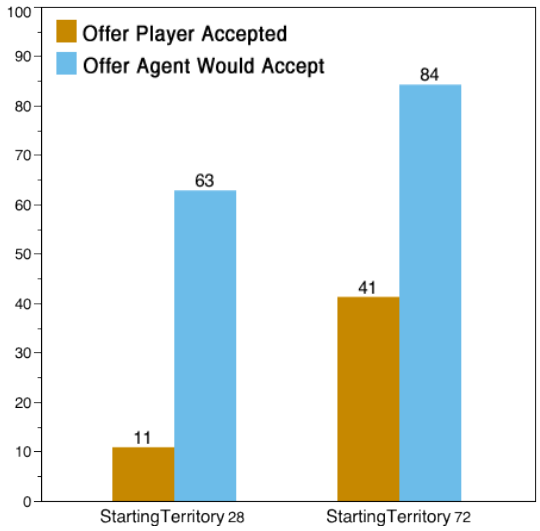


Fig. 2: Offers accepted by participants and agent in Slantchev games.

Similarly, we compared the 145 offers participants accepted (91 from Slantchev28 and 54 from Slantchev72) and the ones the agents would have wanted in that same situation. We model the offers the agent would accept as a uniform distribution across each interval implied by the thresholds in the agent's policy in Section 3.3. Figure 2 shows that, regardless of the starting territory, the offers participants accepted are significantly lower than the ones the Slantchev agent would have wanted (paired-sample t-tests, $p < .0001$ for both pairs). This is consistent with the observation in Section 5.2 that the participants made higher

offers, in that they were willing to achieve a negotiated settlement lower than what the agent would accept. It is also consistent with the hypothesis in Section 5.1 that the participants sought to avoid losing troops in the war more than the agent did, thus lowering their threshold for accepting an offer and bringing about agreement sooner.

6 Impact of Inaccuracies on Agent Performance

Although the participants did not always conform to the agent's model, it is not clear how much the inaccuracies in the agent's theory of mind affected its performance. There are two obvious measures of agent performance: territory owned and troops lost at the end of the game. We use territory as the primary measure because it is a zero-sum game outcome that identifies a "winner" between the agent and participant, whereas both players lose troops in every battle. Given that the agent attempted to maximize territory gains based on its theory of mind about the participants, we might expect its optimal policy to do worse when the participants deviate from those expectations. Indeed, when we examine the agent's performance in the Slantchev conditions, we see that it is negatively correlated with the percentage of participants' actions that deviated from the agent's policy ($r = -.463, p < .0001$). However, we did not observe such a correlation under the Powell conditions ($r = .080, p = .0813$) So while the Slantchev agent earned more territory the more accurate its model of the participant, we cannot make the same overall conclusion under the Powell condition.

6.1 Impact of Deviations in Attacking

Section 5.1 showed that deviations in the participants' attacking decisions occur in only the Powell28 condition, when they reject without attacking. The agent expects them to attack, because the potential gain from winning the war outweighs the 28% territory lost by losing the war. The data confirm this expectation. One-way ANOVA tests show that the agent ended up with more territory when it successfully avoided the war ($Mean_{War} = 57.7\%$, $Mean_{\neg War} = 76.5\%$, $N_{War} = 190$, $N_{\neg War} = 49$, $p = .0026$). Thus, the agent's belief that attacking is the optimal policy for the participants is borne out by the fact that they did worse when deviating, e.g. the agent successfully avoided the war. However, although this divergence between agent expectations and participant behavior works to the agent's benefit, it could do even *better* by holding out for higher offers if it knew that the participant might not attack in retaliation.

6.2 Impact of Deviations in Making Offers

We also analyzed the relationship between the percentage of the participants' "correct" offers and the territorial split at the end of the game. We broke down this analysis based on how the games ended, e.g. agent accepting an offer, winning/losing the war. The rationale is that making a "wrong" offer (e.g. overly

high) may have no immediate impact on the game outcome when it gets rejected, but a "wrong" offer that gets accepted could possibly lead to a much worse outcome for the offeror. Indeed, within the 375 games where the offers were accepted by the agent, there is a negative correlation between the percentage of "correct" offers and territory the participants ended up with ($r = -.3234$, $p < .0001$). This also means that the more often the participants' offers conformed to the agent's policy, the more territory the agent ended up with, as we might expect from a more accurate theory of mind. On the other hand, in the 580 games that did *not* end with the participant making an acceptable offer to the agent (i.e., winning/losing the war, reaching the end of the game, or the participants' accepting an offer from the agent), there is a *positive* correlation between the participants' degree of conformity to the agent's model and their territorial outcome ($r = .3423$, $p < .0001$). In these games, the participants' offers are lower than what the agent wants or expects. Without an agreement that exceeds its threshold, the agent relies on a battlefield outcome or the status quo, which is not likely to favor it as much. This suggests that the agent would benefit by lowering its threshold when dealing with such less generous participants, as the current fixed threshold causes it to miss out on such opportunities.

6.3 Impact of Deviations in Accepting Offers

In the Slantchev games, all of the offers accepted by the participants deviated from the agent's policy, as seen in Section 3.3. To evaluate the impact of this deviation on performance, we compared the 145 games ending with participants accepting the agent's offer and the 333 games ending some other way, e.g. making an offer that the agent accepts or else somehow prolonging the war through all 15 rounds (with both sides keeping their original territory). One-way ANOVA tests show no significant impact on how much territory the participants ended up with ($Mean_{AcceptOffer} = 22.2\%$, $Mean_{Alternative} = 22.5\%$, $p = .9084$). This suggests that even though the participants accepted offers that were too low in the agent's estimation, this deviation had little impact on the game outcome.

7 Discussion

Despite its strong assumptions, the agent's theory of mind still allowed it to perform well (although not optimally), because most of the participants' deviations from expectations did not hurt the agent's performance. By examining the impact of the deviations on the agent's performance, we can prioritize the areas of its theory of mind that we need to improve. For example, by assuming that the participants have complete information, the agent misjudged the offers the participants would make and accept. The agent could instead bias its expectations of the participants to be more conservative given their lack of information. More generally, an agent operating in other domains with similarly uncertain human participants can also benefit from a theory of mind that perturbs an expectation of optimality with a degree of risk aversion.

Even when properly accounting for the participants' uncertainty, we also saw potential deviations due to unanticipated values. For example, the participants' SVI survey responses indicate that they placed a value on reaching an agreement with their negotiation partner. Our agent did not incorporate such a value, nor did it expect the participants to do so. By modifying our agent's theory of mind to include such a value on agreement, it could generate expectations that better account for the participants' willingness to give up territory in exchange for a quicker settlement. We could then repeat the methodology of this paper to evaluate the degree to which the participants' conformed to this modified theory of mind and the potential impact of this change on agent performance.

We can use this paper's methodology to evaluate completely different methods for theory of mind as well. For example, the original game-theoretic analyses of wartime negotiation prescribe that the offering side should start with a low amount and steadily increase it, to screen for the other side's acceptance threshold with minimal over-offering [13, 18]. Even without implementing this model in our own agent, we can still measure the degree to which it matches our participants' behavior. Participants followed this strategy in 317 of our games, and violated it by either repeating or decreasing their offer in 286 games[4]. Overall, there is no significant impact of following the equilibrium strategy on the agent's performance. However, within the games ending with the agent accepting the participant's offer, the ones where the participants played consistently with the equilibrium strategy resulted in more territory for the agent ($Mean_{eq} = 34.6\%$, $Mean_{\neg eq} = 46.9\%$, $p < .0001$). Thus, while screening may be a best response in the theoretical setting, the participants' uncertainty leads them to over-offering in practice, to the agents' benefit.

It is clearly insufficient to evaluate only the accuracy of an IVA's theory of mind with respect to actual human behavior. Some classes of inaccuracy may not have any negative impact on the agent's performance, in which case it is unnecessary to enrich the model to remedy that inaccuracy. Furthermore, in a domain with multiple sources of uncertainty, even one as simplified as our wartime negotiation models, expanding theory of mind without a good model of that uncertainty can be even detrimental to agent performance. Theory of mind exists in service of the overall social interaction, and our analysis demonstrates that we should seek improvements to the modeling of others only when motivated by the subsequent improvements in that interaction.

Acknowledgments

This work was sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). The content does not necessarily reflect the position or the policy of the Government. No official endorsement should be inferred.

[4] The remaining 352 games ended before the participant made a second offer

References

1. Baker, C.L., Tenenbaum, J.B., Saxe, R.R.: Goal inference as inverse planning. In: CogSci. (2007)
2. Curhan, J.R., Elfenbein, H.A., Xu, H.: What do people value when they negotiate? mapping the domain of subjective value in negotiation. Journal of Personality and Social Psychology **91** (2006) 493–512
3. Dupuis, E.C., Cohn, E.S.: A new scale to measure war attitudes: Construction and predictors. Journal of Psychological Arts and Sciences **3**(1) (2006) 6–15
4. Goodie, A.S., Doshi, P., Young, D.L.: Levels of theory-of-mind reasoning in competitive games. Journal of Behavioral Decision Making **25**(1) (2012) 95–108
5. Hoogendoorn, M., Soumokil, J.: Evaluation of virtual agents utilizing theory of mind in a real time action game. In: AAMAS. (2010) 59–66
6. Kaelbling, L.P., Littman, M.L., Cassandra, A.R.: Planning and acting in partially observable stochastic domains. Artificial Intelligence **101** (1998) 99–134
7. Kahneman, D.: Reference points, anchors, norms, and mixed feelings. Organizational behavior and human decision processes **51**(2) (1992) 296–312
8. Kim, J.M., Hill, Jr, R.W., Durlach, P.J., Lane, H.C., Forbell, E., Core, M., Marsella, S., Pynadath, D., Hart, J.: Bilat: A game-based environment for practicing negotiation in a cultural context. International Journal of Artificial Intelligence in Education **19**(3) (2009) 289–308
9. Klatt, J., Marsella, S., Krämer, N.C.: Negotiations in the context of aids prevention: an agent-based model using theory of mind. In: IVA. (2011) 209–215
10. Marsella, S.C., Pynadath, D.V., Read, S.J.: PsychSim: Agent-based modeling of social interactions and influence. In: ICCM. (2004) 243–248
11. Miller, L., Marsella, S., Dey, T., Appleby, P., Christensen, J., Klatt, J., Read, S.: Socially optimized learning in virtual environments (SOLVE). Interactive Storytelling (2011) 182–192
12. Neale, M.A., Bazerman, M.H.: Cognition and rationality in negotiation. Free Press (1991)
13. Powell, R.: Bargaining and learning while fighting. American Journal of Political Science **48**(2) (2004) 344–361
14. Pynadath, D.V., Marsella, S.C.: PsychSim: Modeling theory of mind with decision-theoretic agents. In: IJCAI. (2005) 1181–1186
15. Pynadath, D.V., Marsella, S.C.: Minimal mental models. In: AAAI. (2007) 1038–1046
16. Pynadath, D.V., Marsella, S.C., Wang, N.: Computational models of human behavior in wartime negotiations. In: CogSci. (2013) to appear
17. Robinson, R.J., Lewicki, R.J., Donahue, E.M.: Extending and testing a five factor model of ethical and unethical bargaining tactics: Introducing the sins scale. Journal of Organizational Behavior **21** (2000) 649–664
18. Slantchev, B.L.: The principle of convergence in wartime negotiations. American Political Science Review **97** (2003) 621–632
19. Van Lange, P.A.M., De Bruin, E.M.N., Otten, W., Joireman, J.A.: Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. Journal of Personality and Social Psychology **74**(4) (1997) 733–746
20. Whiten, A., ed.: Natural Theories of Mind. Basil Blackwell, Oxford, UK (1991)