Assessing Common Errors Students Make When Negotiating

Emmanuel Johnson
Institute for Creative Technologies
University of Southern California
Playa Vista, CA
ejohnson@ict.usc.edu

Sarah Roediger
Design Lab
University of Twente
Enschede, Netherlands
s.roediger@student.utwente.nl

Gale Lucas, Jonathan Gratch
Institute for Creative Technologies
University of Southern California
Playa Vista, CA
{lucas,Gratch}@ict.usc.edu

ABSTRACT

Research has shown that virtual agents can be effective tools for teaching negotiation. Virtual agents provide an opportunity for students to practice their negotiation skills which leads to better outcomes. However, these negotiation training agents often lack the ability to understand the errors students make when negotiating, thus limiting their effectiveness as training tools. In this article, we argue that automated opponent-modeling techniques serve as effective methods for diagnosing important negotiation mistakes. To demonstrate this, we analyze a large number of participant traces generated while negotiating with a set of automated opponents. We show that negotiators' performance is closely tied to their understanding of an opponent's preferences. We further show that opponent modeling techniques can diagnose specific errors including: failure to elicit diagnostic information from an opponent, failure to utilize the information that was elicited, and failure to understand the transparency of an opponent. These results show that opponent modeling techniques can be effective methods for diagnosing and potentially correcting crucial negotiation errors.

CCS CONCEPTS

• Computing methodologies~Intelligent agents, Reasoning about belief and knowledge

ACM Reference format:

Emmanuel Johnson, Sarah Roediger, Gale Lucas and Jonathan Gratch. 2019. Assessing Common Errors Students Make When Negotiating. In *Proceedings of ACM Intelligent Virtual Agent Conference (IVA'19), July 2-5, Paris, France*, ACM, New York, New York, USA. 7 pages. https://doi.org/10.1145/3308532.3329470

ACM ISBN 978-1-4503-6672-4/19/07...$15.00.
DOI: http://dx.doi.org/10.1145/3308532.3329470

1 Introduction

Similar to financial literacy, negotiation is an integral skill rarely taught in schools. The US Academy of Sciences and the World Economic Forum identify negotiation as a foundational social skill needed for the future of work through its impact on an organization's creativity and productivity [1] [2]. Deficits in negotiation skills also contribute to the underrepresentation and lack of advancement of women and minorities in STEM fields [3] [4]. Unfortunately, negotiation training is inaccessible to most workers (e.g., even a short 5-day seminar can cost more than $10,000 per student).

Previous research shows that individuals can improve their negotiation skills by practicing with virtual negotiation opponents [5] [6] [7] [8], arguably without eluding the realism that a human negotiator provides. Negotiation requires many skills to reach a favorable outcome. Of particular importance, understanding an opponent's preferences is key to finding hidden agreements, thereby allowing negotiators to create and claim more value [9].

For negotiation training systems to be effective at improving student's negotiation ability, these systems must assess and help students improve their understanding and exploitation of an opponent's wants. In this work, we show how opponent modeling techniques can address this limitation. We also show that these methods are effective regardless of the specific negotiation tactics employed by the automated opponent. We allow participants to practice negotiating with one of a variety of intelligent negotiation agents (each agent implementing a distinct combination of negotiation tactics). We show, in general, participants are bad at understanding their opponent's preferences and this translates into a poor ability to create and claim value. We then show, by analyzing the traces of these negotiations, that opponent modeling techniques can automatically (1) assess the extent to which a negotiator elicited information about their opponent's preferences, (2) quantify how well they utilized the information elicited, and (3) characterize the transparency of their opponent. Such assessments can be used to provide feedback to humans during a negotiation (as in a decision support system), or as part of a personalized feedback system for an intelligent negotiation tutoring system.

The paper is structured as follows. Section 2 introduces the idea of opponent modeling and the benefits of understanding the opponent's perspective. Section 3 highlights our approach in using opponent modeling to assess common errors negotiators make when trying to understand their opponent. In Section 4, we describe our study design and results. Lastly, Section 5 provides a discussion of our results.

2 Opponent Modeling

Generally, opponent modeling can be viewed as the process of trying to understand an opponent's traits and goals from the information gathered through an interaction. Opponent modeling has been applied in a number of adversarial domains from poker playing agent, to agent-agent automated negotiation systems [10] [11] [12] [13]. In this article, we focus on learning the opponent's goals (i.e., the importance they assign to different issues).

2.1 Importance of Opponent Modeling

Opponent modeling is crucial for effective negotiations. It allows negotiators to maximize joint value from a better understanding of their opponent. The more a negotiator knows about an opponent, the better they are at finding win-win solutions. Let us imagine a negotiator is to split a pizza with an opponent. Additionally, let's assume this negotiator only wants the filling and their opponent only wants the crust. A good opponent model would allow a negotiator to integrate an opponent's preferences with their own to propose a win-win deal: i.e., giving their opponent the crust and keeping the rest of the pizza instead of splitting the pizza in half.

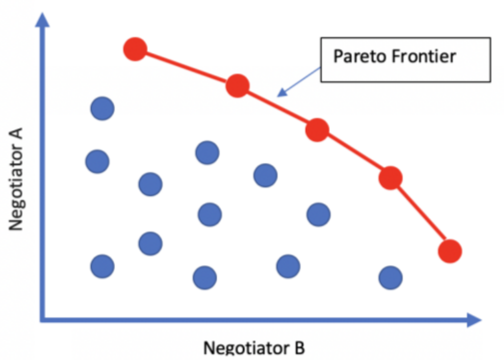

Figure 1: Pareto Frontier

Good opponent modeling ultimately leads negotiators to gain more value for themselves. Integrative deals, such as the one described above, are said to be *efficient*, as they take advantage of tradeoffs across each user's interests. Ideally, negotiators should focus on the *Pareto efficient frontier.* These are deals in which no party can reallocate resources to gain more value without losing value for themselves or their opponent. The pizza deal would be on the Pareto efficient frontier because there is no other way to split the pizza where one person gets more without someone losing a little bit of the items they value. An example of such a frontier is graphically depicted in Figure 1. Deals below this frontier can be improved for one or both negotiators. The problem is, to find the Pareto efficient deals, a negotiator needs to learn their opponent's interest.

2.2 Common Tactic and Errors in Understanding Opponent's Preference

Common tactics for improving one's opponent model involves explicitly asking an opponent about their interest, attending to an opponent's pattern of offers, effectively integrating information shared by an opponent with one's own preferences, and proposing efficient tradeoffs across issues.

Several factors prevent novice negotiators from effectively utilizing these tactics. Negotiators often suffer from a *fixed-pie bias* [14] meaning they assume their opponent has the same preferences as themselves (and therefore, no opportunity for win-win solutions). This bias reduces motivation to learn about the opponent. Research also shows that novices are reluctant to reveal their own information due to fear of exploitation [15].

In contrast, experienced negotiators more freely exchange information. For example, a negotiator may ask explicit questions about their opponent's preference (e.g. "what do you want the most" or "do you like A more than B"). Experienced negotiators also learn about their opponents through their pattern of offers. For example, negotiators tend to claim more of what they want most in a negotiation and concede on less important issues (a process commonly known as logrolling). In doing so, they indirectly communicate their preferences. For instance, in the pizza example from earlier, if the opponent makes an offer to take the crust and leave the filling, this indirectly communicates that they find the crust more important.

Although opponents reveal their preferences through explicit statements or through their pattern of offers, this, by itself, does not allow one to find integrative deals. Novice negotiators often fail to appreciate or utilize the full range of information their opponent provides. To highlight this distinction, we use the term *transparency* to refer to how well an opponent's preferences could be inferred, in theory, from the statements and offers they make during a negotiation. Transparency serves as a theoretically upper bound on the accuracy a human negotiator could have achieved, given the information available to them, although novice negotiators would likely fail to reach this upper bound.

Several factors shape the transparency of an opponent. Some opponent will be more forthcoming about their interests, and thus inherently more transparent. But negotiators can also shape the transparency of their opponent. Experts can enhance transparency by asking the right questions and otherwise drawing out their opponent. In contrast, novices may reduce their opponent's transparency by failing to be transparent themselves (e.g., by withholding information or revealing misleading information).

To summarize, novices tend to make four classes of errors: First, they fail to effectively communicate what they want to their opponent (preventing their opponent from discovering efficient solutions). Second, they fail to elicit information about what their opponent wants. Third, they fail to understand or utilize whatever information their opponent does reveal (preventing themselves from discovering efficient solutions). Lastly, they fail to understand the type of opponent they are negotiating against and their level of transparency. We claim that opponent modeling can help diagnose these four errors.

2.3 Automated Opponent Modeling

Research into automated negotiation agents has yielded effective *opponent modeling* methods for inferring the preferences of an opponent. Later, we will show how to use these techniques to lend insight into the above-mentioned errors. Several techniques have been proposed in the AI literature. These methods differ depending on whether the model generation involves a collection of either offers, preference statement or both.

2.3.1 Modeling From Progression of Offers. Most automated techniques were developed for agent-agent negotiation and attempt to learn from only the pattern of an opponent's offers ([11] provides a good overview of the current state-of-the art). Bayesian and frequency models tend to be the most successful and widely used. Bayesian models try to understand an opponent's preferences by finding the most likely candidate given a set of possible preferences over all issues. They assuming some prior distribution over a set of preference profiles and uses Bayes rule to update their belief given a sequence of observations. Frequency models try to learn weights that represents the relative value of each issue. These models estimate issue weight by noting the frequency with which the value of an issue is offered, as in the N.A.S.H. frequency models [11], or at how often the amount of an issue claimed is changed, as seen in the hardheaded model [16].

2.3.2 Modeling From Offers and Preference Statements. Unlike agent-agent negotiations, human negotiators rely heavily on explicit preference statements [9]. Thus, research on human-human negotiation has extended opponent modeling techniques to integrate this additional channel of evidence. For example, Nazari and colleagues extended the Hardheaded Frequency model to combine information from both explicit preference statements and pattern of offers [17]. (The Hardheaded model was selected because it proved to be the most accurate model in the 2011 Autonomous Negotiating Agent Competition). To contrast the value of these different information channels, Nazari proposed three models: a model based solely on the pattern of offers, a model based solely on explicit preference statements, and a model that utilized both channels:

Offer-Only Model: Following the Hardhead Frequency model, this model makes two assumptions. First, if an item is valuable to a negotiator, they will claim more of it for themselves. Second, if an item is valuable to a negotiator, they will claim that item more frequently. Thus, if an issue (i) is discussed in an offer (k), it tells you how much of that issue was claimed for self (l_k) and how much level was assigned to the opponent (l'_k). Thus, to calculate the weight for each item we compute a ratio of the items claimed for self over items given to opponent

$$w_i = \frac{l_k}{l'_k}$$

Statement-Only Model: This model estimates weights from the explicit preference statements by counting how often a positive or negative statement is made about an item. For example, "I like gold more than spices" is a positive statement about gold whereas "I like gold less than bananas" is a negative statement towards gold. The weight for each item i is computed as follows:

$$w_i = |P_i| - |N_i|$$

where Pi is a count of all positive statements made about an item *and* N_i is a set of all negative statements made about an item.

Offer-Statement Model: This model integrates the two information channels by averaging the weights of the Offer-Only and Offer-Statement models for each item.

3 Using Opponent Modeling for Assessment

The goal of our work is to demonstrate well-established human automated modeling techniques provide novel insight into common errors individuals make when negotiating with a pedagogical agent. As discussed in Section 2.2, novices make several errors: they fail to communicate their preferences to their opponent, they fail to elicit their opponent's preferences, and they fail to utilize the information available to them.

In contrast to human negotiators, automated methods seem quite accurate at inferring the opponent's model. Nazari [18] showed that the Offer-Statement model was highly accurate and outperformed models that focused solely on offers or preference statements. This suggests that, for a given negotiation, the Offer-Statement model could serve as an approximate measure of the transparency of the opponent (recall, as discussed in Section 2.2, transparency is how well an opponent's preferences are inferable. It serves as a theoretically upper bound on the modeling-accuracy a human negotiator could have achieved, given the information available to them).

If we assume the Offer-Statement model is a good approximation of the transparency of a negotiator, several conclusions naturally follow (which we empirically test in Section 4). First, if the Offer-Statement model fails to learn an accurate model of a player, this implies errors in information exchange: If humans show low transparency, this suggests they erred in providing useful information to their agent opponent; If the agent is opaque, this implies the human negotiator erred by failing to elicit useful information from the agent (though different types of opponents may also differ in their propensity to share). Second, if the Offer-Statement model is more accurate than the human at predicting the agent's preference, we can infer that the human erred in utilizing the information available to them. For example, they may have failed to attend to one of the channels of information that the Offer-Statement model uses. Third, by analyzing differences in the three opponent models (Offer-Only, Statement-Only, and Offer-Statement), we can further diagnose which channel the human negotiator likely ignored. For example, if an opponent's Offer-Only model is more accurate than the Statement-Only, this implies that the information was mainly present in the pattern of offers. Thus, if the human's accuracy was low, they likely failed to understand the information contained within these offers.

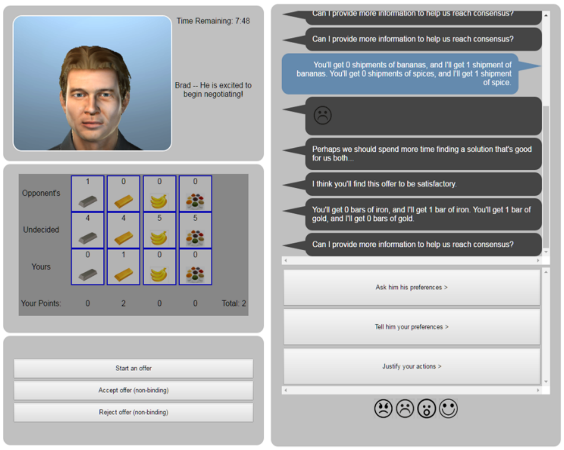

Figure 2: IAGO Agent

3.1. Classifying Novice Negotiators by Error

Opponent models provide a continuous measure of transparency, but in assessment it is often useful to discretely classify students into separate categories such as "doing well" versus "needs work" for the purposes of feedback. In the remainder of the paper we divide participants into tertiles (three equal-sized groups) based on how well they understood their opponent's preferences: A-students, B-students, or C-students. A-students can essentially be seen as experts whereas the other groups should be targeted for feedback. In the experiments that follow, although we correlate continuous modeling-accuracy with outcomes, we focus our analysis on these discrete grouping as they help to better visualize the consequences of failing to model one's opponent.

4. Assessing Student's Modeling Abilities

To validate our approach, we recruited a panel of participants to practice negotiating against one of several possible automated opponents. We assessed how well students understood their opponents and how this impacted their ability to create and claim value. We then tested if our proposed method yields the predicted insights into student errors. All participants negotiated with the IAGO platform (see figure 2) [19]. IAGO is an online human-agent negotiation platform that allows developers to build virtual human agents that negotiate with a human user. Within IAGO, negotiators can exchange offers and exchange explicit preference information (do you like A more than B?) with the agent. It has been used by a number of researchers to build human-like negotiation agents that employ a variety of common negotiation tactics [20].

In piloting our models on human-agent data we discovered one difference: people tend to exchange fewer offers and information with IAGO agents than what Nazari found in her human-human corpus. With IAGO, negotiators exchanged on average 3.7 offers and 2.08 preference statements and the agent exchanged 3.11 offers and 3.43 preference statements. In [18] corpus, humans exchanged on average 5.8 offers and 9.9 preference statements. The consequence is that the model often fails to recommend a difference between issues. This led us to make some small adjustments to Nazari's method (described below)

In order to evaluate a negotiator's ability to understand an opponent's wants, models must assess their errors in following expert negotiation principles. In this section we focus on diagnosing errors. We describe our opponent models' ability to understand the transparency of a human user, their ability to shape the transparency of their opponent (e.g., by eliciting information), as well as their ability to utilize the information provided by an opponent.

4.1 Study Design

Opponent: Participants engaged in a single multi-issue negotiation task with an IAGO agent. To ensure that the results were not specific to the behavior of the automated opponent, participants negotiated against one of four possible agents. Agents varied in terms of two common differences found amongst human negotiators. First, agents varied as to whether or not they used *anchoring* [21]. Anchoring is a negotiation tactic that involves making a very strong initial offer and has been found to help negotiators claim more value. Second, agents varied as to whether they adopted a "fixed-pie bias" [17]. When negotiators exhibit a fixed-pie bias, they approach a negotiation with the assumption that their opponent wants the same things as they do unless the opponent reveals information that contradicts this assumption. Non-fixed-pie agents assume the negotiation is maximally integrative unless the opponent revealed otherwise. Other than these two factors, the agents followed the default "Pinocchio" agent behavior provided by the IAGO agent platform [22]. Anchoring and bias were manipulated independently to yield four agent types.

Participants: A total of 609 participants who were English speakers from America were recruited via Prolific, an online platform similar to Amazon Mechanical Turk which is often used for recruiting research participant. Of the 609, 132 were removed for failing to pass the attention check questions and other requirements, leaving 477 negotiations in the corpus. To motivate their performance, participants were paid for their participation in the study and provided tickets into a $100 lottery proportional to their outcome in points. This study required participants to be native English speakers from America because of the language used so as to ensure participants understood the instructions and agent preference information.

Negotiation Task: Participants engaged in a multi-issue bargaining task in which they and the agent had to divide a number of items amongst each other. The negotiation took a total of 10 minutes. The items to be divided are as follows; 7 bars of gold, 5 bars of iron, 5 shipments of spices, and 5 shipments of bananas.

Table 1: Agent and Human Payoff Matrix

	Gold	Iron	Spices	Bananas
Agent	4	1	2	3
Human	4	3	2	1

Both the agent and participant had unique preferences across the items, and neither the agent or the participant knew the other's preference. The payoff metric for each negotiator is shown in Table 1. Prior to the negotiation participants were told how much each item was worth to them. In addition to the worth of items, participants were also told they would receive only four points if they failed to reach an agreement. The task allows the opportunity to create value. Agreements can be made more efficient by trading off value between iron and bananas. The joint value of the final deal is maximized if the participant claims all the iron and the agent claims all the bananas. Gold and spices are fixed-pie issues. Participants can create more efficient solutions if they correctly model their opponent's preferences.

Measures: We extract several measures of negotiation performance from the IAGO negotiation logs:

Outcome measures: To assess the quality of the outcome, we measure the individual points obtained by the participant and the joint points (i.e., the sum of individual points obtained by the participant and the agent). Participant points is a measure of value claiming. Joint points is a measure of value creation.

Opponent modeling measures: To assess opponent models, we collect four measures. Following the negotiation, we ask participants to rank the priorities of their opponent to give insight into how well they understood their opponent's preferences. We then ran the three automatic models (statement model, offer model, and dual model) over the IAGO logs to give an "expert" opinion on how well the opponent could have been modeled, in principle, from the various information channels. Each of these approaches yields a ranking over the opponent's priorities. We then adopt a standard approach to quantify the accuracy of these four models.

A number of approaches have been proposed for assessing the accuracy of an opponent model. Baarslag and colleagues [11] provides an overview of the state-of-the-art in evaluating opponent modeling technique. One common measure used is assessing the accuracy of an opponent model is the rank distance. Given that it's a common practice to represent agent's preference as a rank ordering over a set of issues, we felt that it would be the best metric for measuring differences between ranking. This is done by comparing the utility of all possible deals (Ω) in the outcome given a rank r_a and rank r_b, and computing the average number of conflicts:

$$d_r\left(r_a, r_b\right) = \frac{1}{|\Omega|^2} \sum_{\omega\in\Omega,\omega'\in\Omega} c < r, < r'(\omega, \omega')$$

As mentioned above, we made minor changes to adapt Nazari's Offer-Only models due to the low level of offer exchange found with IAGO agents (the statement model and Offer-Statement models remained unchanged). In the Offer-Only model, Nazari computes a ratio of items a negotiator claims over the items that are allocated to an opponent. One source of information which is ignored is the items left on the table (items not claimed by either party). In our model, we treat items left on the table as items the negotiator does not want. So instead of computing the weight per issue as a ratio of items claimed over items given to opponent, we incorporate the information about items left on the table. The weight for each item is updated as follows, where l_k is the items claimed by a negotiator, l'_k is the items given to an opponent and l_u is the items left unclaimed.

$$w_i = l_k\,(l'_k + l_u)$$

As prior research indicates that people tend to assume their opponent wants the same things as them in the absence of information (the fixed-pie bias), to help resolve ambiguity in the model and incorporate more human-like bias, we break ties in issue weights by generating a set of all possible ranking given current knowledge. We then compute the rank distance between all possible rankings and the fixed-pie bias (negotiator's preferences). The closest preference rankings to the fixed-pie bias is the rank which is selected. These changes yielded higher accuracy on our pilot data.

4.2 Results

4.2.1 Quality of Negotiated Agreements. We have claimed that novice negotiators have difficulty creating and claiming value because they fail to understand their opponent's preferences. To verify this hypothesis, we examined the correlation between the participant's accuracy in modeling their opponent and their individual and joint points. We confirm that participant accuracy is highly correlated with both joint (r = .445, n = 479, p <.001) and individual points (r = .295, n = 479, p <.001). In other words, if participants failed to understand their opponent, their deals were inefficient, and they claimed less points for themselves.

We performed a mediation analysis to understand if participants' inability to claim value was solely due to their failure to create value or if poor opponent modeling has wider negative consequences. We find that joint points partially mediate the impact of modeling accuracy on individual points. We see a highly significant direct relationship of model accuracy to user points ($\beta = .295, t(1,477) = 6.743, p < .001$). There is also a highly significant relationship between model accuracy and joint points ($\beta = .366, t(1,477) = 8.598, p < .001$). However, when both

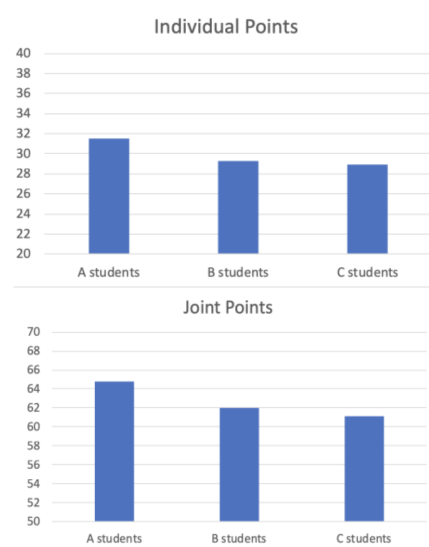

Figure 3: Value Claimed and Created by Group

independent variables and mediator are examined, the significance of the model accuracy to user points decreases from p < .001 to p=.002. The Sobel test statistic = 5.579, p < .001. In other words, participants claim less value, in part because they fail to grow the pie, but also because of their inability to understand an opponent generally undermines their earnings.

To test the benefit of groups participants into A-, B-, and C-Students (as discussed in Section 3.1), we perform a one-way ANOVA to analyze individual and joint points across groups to see if they show qualitative differences in negotiated outcomes. As expected from the above mentioned correlations, individual points differed significantly across groups (F(2,474) = 19.709, p < .001). In pair-wise comparisons, A-students (i.e., those that most accurately inferred their opponents preferences) earned significant more individual points than B-students (t(2,474) = -4.618, p < .001) and C-students (F(2,474) = -5.993, p < .001). There was not a significant difference between the B and C students (F(2,474) = -1.375, p < .170). See Figure 3.

Similarly, we see a difference across groups in terms of joint value (F (2,474) = 44.918, p < .001). As with individual points, participants in the A group, generated significantly more joint points than those in the B group (t(2,474) = -6.895, p < .001) and C group (t (2,474) = -9.092, p < .001). We also see a significant difference in individual points between the B and C student groups (t (2,474) = -2.227, p < .026). See Figure 3. Together, this shows that grouping students in terms of their opponent modeling accuracy yields considerable differences in individual and joint value. This provides empirical support for the utility of our A/B/C-student classification.

4.2.2 Understanding of Negotiator's transparency. We claimed that opponent models could serve as an objective way to characterize how transparently a negotiator communicates their preferences and which channel (statement vs. offers) is most diagnostic. Here we examine if this notion of transparency gives insight into the behavior of the agent and student negotiators.

Agent Transparency: The different agents adopt quite different tactics and we expect this should impact their transparency. To test, this, we examined the transparency of the different automated agents based on their type (anchoring and fixed-pie bias). Figure 4 shows the accuracy of the three automated models and the users' estimate broken out by the four agent types: optimistic (i.e., no fixed-pie) anchoring, optimistic no-anchoring, fixed-pie anchoring, and fixed-pie no-anchoring. This Figure also shows the result of collapsing across agent type (average agent).

Overall, as expected, the most accurate inferences come from combining both information channels (i.e., the Offer-Statement model), though the statement model also yields reasonable accuracy. The offer only model performed worse across most agents. Users performed uniformly poorly in their estimates.

The results also show differences in which channel was most diagnostic depending on the agent type. There were significant, sizable differences between the different models (F(3,1425) = 169.66, p < .001; see average agent in Figure 4), and the models also varied in their differences by agent type (F(3,1425) = 28.43, p < .001; see remainder of Figure 4).[1] The differences are driven by the optimistic-anchoring agent. This can be explained by the fact that, as this agent assumes there is a win-win solution, it leads with a strong initial offer that incorporates tradeoffs (it claims all of what it wants most while offering the participant all of what it wants least). Note also that the fixed-pie agents are the least transparent when it comes to their offers. Again, this can be explained by the fact that, in contrast to optimistic agents, fixed-pie agents split issues evenly unless the participant reveals their own asymmetric preferences. Thus, the offers of fixed-pie agents provide little information about their true preferences (just as tends to occur in human negotiators that hold this bias)

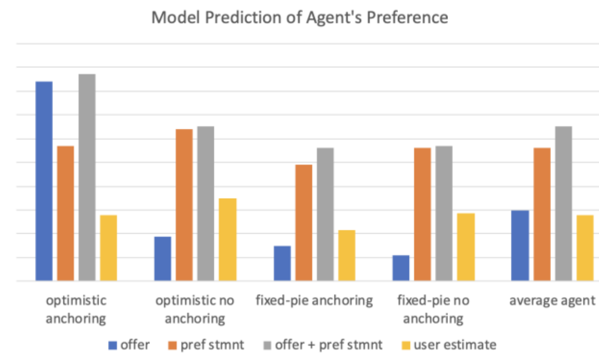

Figure 4: Accuracy of Model by Agent Type

Participant Transparency: Opponent models should be able to provide insight into how well they are communicating their own preference information. To test this, we examined the transparency of the human participants by group (see Figure 5). We see clear differences in transparency by group, with A-students the most transparent (F(2,474) =11.762, p < .001) . We also see that the Offer-Statement model is better at predicting the preferences of A-students, though these are less transparent than the automated agents (F(2, 474)=156.860, p < .001). Unlike the automated agents, participants communicate more information through their pattern of offers. This suggests that even A-students fail to communicate their preferences to their automated opponent. This means their opponent will have difficulty helping them to create value.

[1] Because the agents differed from each other on two dimensions – belief (optimistic vs. fixed pie belief) and anchoring (anchoring vs. no anchoring), we conducted a 2 (belief) X 2 (anchoring) X 4 (model) mixed ANOVA, where the first two factors are between-subjects and the latter is within-subjects. All other effects were significant at p > .001, but they are not reported here because they are all qualified by the 3-way interaction, which we reported (and depicted in Figure 4).

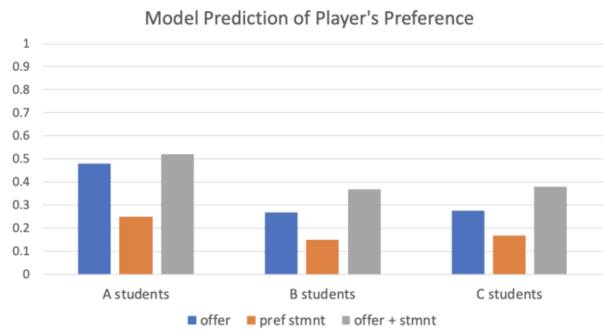

Figure 5: Models' Insight into Student's Preference

4.2.3 Investigative skills of a Negotiator. The above analysis shows that some negotiators are easier "to read" than others. While this is partially due to the characteristics of the negotiator, it also reflects their opponent's skill in drawing out diagnostic information through asking good preference questions and through exploring tradeoffs in their pattern of offers. We next examined how well students could draw out diagnostic information from their automated opponent. To do this, we examined how accurately we could infer the automated agent's preferences based on the skill of the human participant (i.e., A-students vs. B-students vs. C-students). If one group is better at interrogating their opponents, this should allow the automated techniques to more accurately predict these opponent's preferences. Figure 6 shows how accurate different opponent models were at predicting the agent's preference compare to the user broken out by groups. A students are as good as the Offer-Statement model at estimating their opponent's preferences ($F(1,316) = 1.747$, $p < .5$). This suggests that A students are effective at integrating both offer and preference information into their estimates. The B and C students performed much worse than the agent-based models.

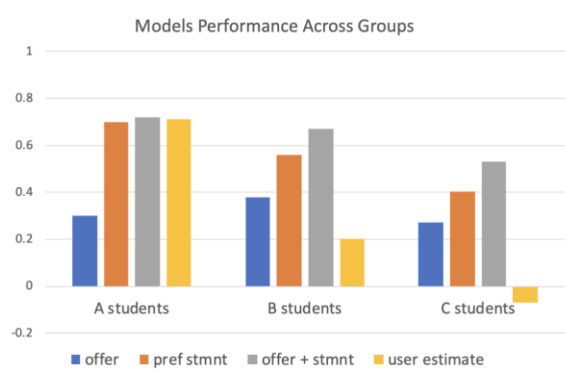

Figure 6: Accuracy of Agent's preference across Groups

4.2.4 Information Utilization. Finally, just because a student can draw out diagnostic information from their opponent, this doesn't mean they are effective at combining this information into an accurate model of their opponent. We examined if the agent-based models can give insight into the type of errors that participants are making when given the appropriate information. We ran the agent-based models on each group separately and compare these results with the participants' estimates.

These results in Figure 6 illustrate several points. First, when comparing the automated opponent model (Offer-Statement) to the user's estimated model, we see a significant drop in accuracy for the B-students ($F(1,316) = 39.032$, $p < .001$) and C-students ($F(1,316) = 161.342$, $p < .001$). This illustrates that the information was available to novice negotiators, but they failed to properly attend to this information. Figure 6 also illustrates another important difference between users. Note that, although the B- and C-students failed to exploit the information available to them, the A-students performed just as accurately as the Offer-Statement model. This indicates that expert negotiators are not simply better at using the preference information available to them, but they are better at eliciting diagnostic information from their opponent.

5 Discussion and Conclusion

Opponent models are a useful tool for assessing common errors that negotiator's make in creating value. They provide a diagnostic tool for understanding how well one understands their opponent's preference, as well as their ability to use that information to find a win-win solution. In this work we showed that although the human-agent negotiation is different from human-human and agent-agent negotiations, opponent modeling techniques from both domains can be used in human-agent negotiation to asses negotiator's value creation ability. To evaluate the success of our models, we examined their performance in relations to the negotiator's ability to infer the agent's preferences across various agent types. We showed that our models were good at diagnosing both the agent and human preference modeling ability. Using our models, we can also determine how much information the agents and participant are revealing about themselves vs how much they are gaining from their opponent. This information can be used to determine if the participant is providing too much information about themselves which could lead to exploitation from their opponent, or if the participant is making effective use of the information the agent shared. Here our focus is on interactions with a virtual pedagogical agent, although our proposal should equally provide insight into human-human negotiations (assuming the appropriate annotations are available).In the future we will extend this work by incorporating our opponent models into a negotiation training agent's feedback system. We also plan to examine other types of models for understanding what an opponent wants. In this work we use very simple frequency models. However, there are other more complex models such as the Bayesian models mention earlier and neural network approaches that may become even more fruitful for negotiation training.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation Grant 1822876 and the US Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

[1] National Academies of Sciences, Medicine and others, Promising practices for strengthening the regional STEM workforce development ecosystem, National Academies Press, 2016.

[2] W. E. Forum, "The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution," in Global Challenge Insight Report, World Economic Forum, Geneva, 2016.

[3] E. G. Goldman, "Lipstick and labcoats: Undergraduate women's gender negotiation in STEM fields," NASPA Journal About Women in Higher Education, vol. 5, pp. 115-140, 2012.

[4] M. Hernandez and D. R. Avery, "Getting the Short End of the Stick: Racial Bias in Salary Negotiations".

[5] J. Broekens, M. Harbers, W.-P. Brinkman, C. M. Jonker, K. Van den Bosch and J.-J. Meyer, "Virtual Reality Negotiation Training Increases Negotiation Knowledge and Skill," in Intelligent Virtual Agents: 12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012. Proceedings, Y. Nakano, M. Neff, A. Paiva and M. Walker, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 218-230.

[6] R. Lin, Y. Oshrat and S. Kraus, "Investigating the benefits of automated negotiations in enhancing people's negotiation skills," in Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1, 2009.

[7] S. Monahan, E. Johnson, G. Lucas, J. Finch and J. Gratch, "Autonomous Agent that Provides Automated Feedback Improves Negotiation Skills," in International Conference on Artificial Intelligence in Education, 2018.

[8] A. Rosenfeld, I. Zuckerman, E. Segal-Halevi, O. Drein and S. Kraus, "NegoChat: a chat-based negotiation agent," in Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems, 2014.

[9] L. L. Thompson, "Information exchange in negotiation," Journal of Experimental Social Psychology, vol. 27, pp. 161-179, 1991.

[10] K. Laviers, G. Sukthankar, D. W. Aha, M. Molineaux, C. Darken and others, "Improving Offensive Performance Through Opponent Modeling.," in AIIDE, 2009.

[11] T. Baarslag, M. Hendrikx, K. Hindriks and C. Jonker, "Predicting the performance of opponent models in automated negotiation," in Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02, 2013.

[12] D. V. Pynadath and S. C. Marsella, "PsychSim: Modeling theory of mind with decision-theoretic agents," in IJCAI, 2005.

[13] J. Klatt, S. Marsella and N. C. Krämer, "Negotiations in the context of AIDS prevention: an agent-based model using theory of mind," in International Workshop on Intelligent Virtual Agents, 2011.

[14] C. K. W. De Dreu, S. L. Koole and W. Steinel, "Unfixing the fixed pie: a motivated information-processing approach to integrative negotiation.," Journal of personality and social psychology, vol. 79, p. 975, 2000.

[15] R. M. Coehoorn and N. R. Jennings, "Learning on opponent's preferences to make effective multi-issue negotiation trade-offs," in Proceedings of the 6th international conference on Electronic commerce, 2004.

[16] T. Ito, M. Zhang, V. Robu and T. Matsuo, Complex automated negotiations: Theories, models, and software competitions, Springer, 2013.

[17] Z. Nazari, G. Lucas and J. Gratch, "Fixed-pie Lie in Action," in International Conference on Intelligent Virtual Agents, 2017.

[18] Z. Nazari, G. M. Lucas and J. Gratch, "Opponent modeling for virtual human negotiators," in International Conference on Intelligent Virtual Agents, 2015.

[19] J. Mell and J. Gratch, "IAGO: Interactive Arbitration Guide Online," in Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, 2016.

[20] J. Mell, J. Gratch, T. Baarslag, R. Aydoğran and C. M. Jonker, "Results of the First Annual Human-Agent League of the Automated Negotiating Agents Competition," in Proceedings of the 18th International Conference on Intelligent Virtual Agents, New York, NY, USA, 2018.

[21] A. D. Galinsky and T. Mussweiler, "First offers as anchors: the role of perspective-taking and negotiator focus.," Journal of personality and social psychology, vol. 81, p. 657, 2001.

[22] J. Mell and J. Gratch, "Grumpy & Pinocchio: Answering Human-Agent Negotiation Questions Through Realistic Agent Design," in Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, Richland, 2017