Fixed-pie Lie in Action

Zahra Nazari ✉, Gale Lucas, Jonathan Gratch
University of Southern California,
Institute for Creative Technologies, Los Angeles, California
Email: znazari@usc.edu, {lucas, gratch}@ict.usc.edu

Abstract. Negotiation is a crucial skill for socially intelligent agents. Sometimes negotiators lie to gain advantage. In particular, they can claim that they want the same thing as their opponents (i.e., use a "fixed-pie lie") to gain an advantage while appearing fair. The current work is the first attempt to examine effectiveness of this strategy when used by agents against humans in realistic negotiation settings. Using the IAGO platform, we show that the exploitative agent indeed wins more points while appearing fair and honest to its opponent. In a second study, we investigated how far the exploitative agents could push for more gain and examined their effect on people's behavior. This study shows that even though exploitative agents gained high value in short-term, their long-term success remains questioned as they left their opponents unhappy and unsatisfied.

Keywords: Human-Agent Negotiation, Behavioral Game Theory, Deception

1 Introduction

Negotiations are an essential part of our everyday social interactions. Any socially intelligent entity must consider negotiation as an important domain to perform successful interactions with humans. Virtual humans with negotiation capabilities can support novel forms of electronic commerce [1], teach people how to improve their negotiation skills [2][3][4] and, more generally, serve as a challenge problem to advance intelligent virtual agent research [5].

One of the key social skills in human negotiations is how to exchange information about what each side wants. Unlike the rational analysis of negotiation in game theory [6] (where such communication is often considered "cheap talk" that can't be verified and thus be ignored), or most computational approaches to negotiation (where agents simply exchange offers [7] and must indirectly infer what the other side wants), human negotiators talk directly about their goals and interests in a negotiation. In general, information exchange helps human negotiators better understand each side's perspective, often leading to better deals on both sides [8]. Unfortunately, this exchange of information provides an opportunity for malicious actors to use deception for gaining disproportionate advantage.

Research has recently explored the potential of agent technology to understand and model how people use deceptive information exchange to maximize their own self-interest. For example, Gratch and colleagues [9] introduced the *misrepresentation* *game*, a game theoretic analysis of how a deceptive party could misrepresent its preferences to gain benefit, while seeming fair and honest. They proved a particular communication strategy, called the *fixed pie lie*, is the optimal solution to the misrepresentation game and provided some evidence that the technique was effective against human participants. Their proof and evidence, however, rely on assumptions that are unlikely to hold in real-world situations. The liar asks questions about preferences of the user in a one-way information exchange and ends with an ultimatum-type take-it-or-leave-it complete offer. In more realistic negotiations, however, parties are free to ask any questions they like, make tentative offers, and make counter-proposals to offers they receive. Also, both parties are free to lie. Thus, it is unclear if an agent can use this strategy in more realistic settings. In this paper, we explore the misrepresentation game in a more realistic context. Building on a platform that supports human-agent negotiation [10], we investigate Gratch and colleagues' solution to the misrepresentation game under far less restrictive conditions. We show that the fixed-pie lie is surprisingly effective in practice. These results give insight into how people use deceptive communication. They also highlight ways that people (and potentially automated agents) can exploit naïve negotiators. Ultimately, such models facilitate techniques to recognize and protect against such practices.

2 Misrepresentation Game

Gratch and colleagues [9] introduced the *misrepresentation game* to analyze how a deceptive actor could best lie to gain advantage. In negotiations, parties try to decide over several issues under discussion. Each of the parties have their own priorities and preferences, and they should consider the interests of their opponent to come up with an agreement that is acceptable to both. Negotiators, however, are usually not aware of their opponent's preferences and must infer them through exchange of information and offers. Discovering preferences of the opponent is a challenge for both human and automated negotiators [11], [12]. This information exchange provides the opportunity for malicious actors to misrepresent their own preferences to gain advantage. The misrepresentation game uses a game theoretic analyses of how a deceptive agent could best lie, and provides an optimization solution to this problem. Here, we briefly describe this game.

The *misrepresentation game* is defined in the following model of negotiation, known as bilateral multi-issue bargaining. There are two parties that must reach agreement over a set of issues. Each issue can have one of many possible levels, and the value that each party derives from the agreement can be written as a utility function over the assigned levels for the issues. In the misrepresentation game [9], the party under consideration (called the *deceiver*) tries to claim a false set of preferences such that a personally-advantageous solution appears as a fair and efficient solution to the opponent. More formally, let a_i be the deceiver's true preferences for issue *i,* b_i be the true preference of the opponent for this issue, and $\mathrm{x_i}$ and $\mathrm{y_i}$ be the agreed upon levels for these players. The objective is to find a 'false' preference $\overline{a}_i$ that maximizes the deceiver's utility, given that the negotiated solution appears to be efficient and fair if $\bar{a}_i$ is believed. This can be written as the following optimization problem (under the assumption that issues are independent).

$$\max_{\bar{a}_i} \sum_{i=1}^{n} a_i \,.\, x_i \qquad (1)$$

where x_i is the optimal solution of

$$\max_{x_t, y_t} \bar{a}_i \,.\, x_i \qquad (2)$$

such that :

$$\sum_{i=1}^{n} \bar{a}_i \,.\, x_i \leq \sum_{i=1}^{n} b_i \,.\, y_i \qquad (3)$$

$$\forall i : x_i + y_i \leq L_i \qquad (4)$$

$$0 \leq x_i, y_i \leq L_i \qquad (5)$$

Here L_i is the total number of levels (discrete case) or amount of resource (continuous case). The optimization program (2-5) capture the problem of finding the optimal solution under the fairness constraint (3), known as Kalai's fairness. It can be shown that this optimization problem can be solved by mixed integer-linear programming [13].

Gratch and colleagues [9] showed that the deceiver's best strategy in the misrepresentation game is to always pretend that the negotiation has a distributive structure (i.e., the two parties have the same valuation on issues), even if it does not. That is, honesty is the best policy in distributive negotiations, but in an integrative negotiation a malicious player can gain advantage from a "fixed-pie lie." Either way, a malicious player should simply pretend the same preferences as their opponent's. After building a sufficiently complete model of the opponent, the deceiver then proposes a deal that gives the opponent a disproportionate amount of the opponent's favorite issue, and keeps a disproportionate amount of the rest. This deal appears fair but benefits the deceiver.

Gratch and colleagues [9] provided evidence for effectiveness of this strategy in a simplified setting. In their setting, the agent started by asking questions about preferences of the user in a one-way information exchange and ended with an ultimatum type take-it-or-leave-it complete offer. In more realistic negotiations, however, parties are free to ask any questions they like, make tentative offers, and make counter-proposals to offers they receive. Also, both parties are free to lie. Thus, it is unclear if an agent can use this strategy in more realistic settings.

In this paper, we aim to answer this question using a publicly available platform called IAGO [10] that provides a realistic environment for negotiations between agents and humans. A few characteristics that make IAGO a great choice for our purpose are: 1) It supports partial offers, 2) Offer exchange is not limited to alternative offer protocol, 3) Visual representation of emotional signals are provided and 4) Communication of preferences are made possible. This platform is used in human-agent competitions in the IJCAI conference and is a powerful framework to conduct human-agent negotiations.

3 Framework and Implementation

The goal of this work is to investigate if virtual agents can successfully use fixed-pie lie strategy in realistic negotiation settings. We use IAGO which is the only agent-human negotiation platform that provides many characteristics of natural human-human negotiation settings. In this section first we briefly describe the IAGO platform, and next we explain the architecture of our agent that was used to implement and test the fixed-pie lie strategy.

3.1 IAGO Platform

IAGO is a web-based platform proposed for conducting negotiations between agents and humans. It consists of a user interface that human users can interact with and an API that agent designers can use for their agent implementation.

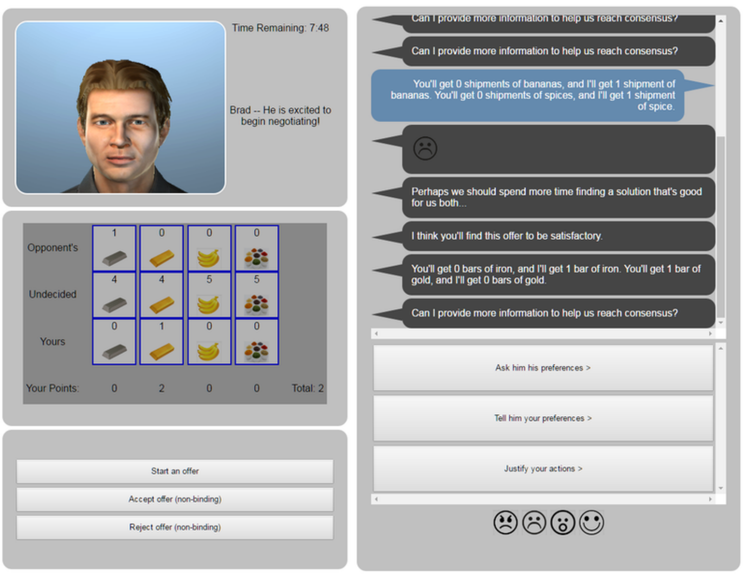

Fig. 1. IAGO's User Interface

IAGO provides a web-based interface that could be displayed on a web browser via a link. The user then can interact with the implemented agent using the environment shown in Fig. 1. The window on right shows the conversation between the agent and the user. Negotiation partners can exchange full or partial offers (e.g., I'll offer you 1 bars of iron and take 3), ask and exchange information about preferences (e.g., "what do you like best? I like iron more than copper") and exchange emoji' that to convey emotional state.

3.2 Agent design

To successfully apply and examine “fixed-pie lie” strategy and compare it to control conditions, our agents must be able to handle three critical requirements: 1) They must make an accurate model of the opponent’s preference profile (User Modeling), 2) communicate an intended preference profile (User Influence) and 3) evaluate the utility of deals to come up with desired offers, or accept/reject received offers (Decision Making). In this section, we briefly present the mechanisms by which our agents implement each of these three components.

User Modelling. The agent starts with no information about users’ preference profile and attempts to acquire maximal information through communication with them. The agent explains to the user about the importance of communication and asks questions about user’s preferences. At each point in the negotiation the agent saves a space of possible preferences for the user based on the information it has gathered up to that point. For example, in the beginning and for a three issues negotiation over bars of gold, silver, and iron, this space is constructed of six possible preferences (Table 1.a). If the user expresses that she likes gold best, then only two possible preference profiles remain (Table 2.b), and after user’s answer to if she likes silver or iron more (Table 1.c for iron), the agent has a solid estimation of the user preference profile.

Table 1. tables show the space of possible utility weights for all issues under negotiation.

Gold	Silver	Iron
1	2	3
1	3	2
2	1	3
2	3	1
3	1	2
3	2	1

Gold	Silver	Iron
3	1	2
3	2	1

Gold	Silver	Iron
3	1	2

1.a 1.b 1.c

When asked about preferences, users may provide contradictory information. This could be due to laziness, lying or just unintentional errors. To detect such incidents and keep the user expressions consistent, we implemented a “contradiction check” element that is called after each preference expression made by the user. If the new expression was consistent with the previous ones, the knowledge base gets updated; otherwise, the user is informed about the inconsistency and asked for correction. The following is a description of how contradiction check works in our agents.

Contradiction Check.

To verify that a user’s preference expression is consistent with her previous statements, the agent keeps a graph representation of preferences expressed by the user. Each node in this graph represents an issue, and each edge represents a relative connection. The method will add a new expression (and the corresponding edges) only if it does not create a loop in the graph. Otherwise, the contradiction check fails and the user is notified. For example, in a negotiation over three issues of gold, iron, and silver, the graph has three nodes, one for each issue. If a user states “I like gold the best”, edges are is added from all other issues to gold (here, iron and silver). If the user then expresses “I like silver more than gold”, adding this new edge could create a loop in the graph (from gold to silver, and silver to gold); The agent will inform the user about the inconsistency and asks for correction.

Shaping the User's Model of the Agent. During the negotiation, the agent provides the user with information about the agent’s preferences. This information could be in the form of voluntary information exchange in return for users’ similar behavior, or by answering questions asked by the user. A variable “Honesty” specifies how to shape the user’s model of the agent. A truthful agent will communicate its actual preference profile; however, a liar agent would communicate the same preference profile as the user.

Decision Making. Once an agent has a proper model of the user’s preferences, along with its own preference profile, the agent can calculate how much each deal is worth for itself and how much for the user. The percentage of the joint outcome value that the agent receives is the criterion evaluating each deal. For example, if a deal is worth 22 points for the agent and 22 points for the user, a fair agent will accept it, but an exploitative one will reject it. The same module is used to pick an offer to propose from the agent side.

3.3 Negotiation Task

The misrepresentation game illustrates how a negotiator can use the integrative potential in a multi-issue negotiation task to gain advantage while seeming fair. Even though this “fixed-pie lie” strategy works best when the task is fully integrative, we use a partial integrative task to demonstrate the generalizability of the strategy[1]. The negotiation was framed as a “Resource Exchange Game” where opponents negotiated over four issues of gold, silver, copper and iron. In this task, negotiators should distribute four bars from each issue (overall 16 bars) between themselves. The preference profiles for the agent and the user (shown in Table 2) are used as weights to calculate the gained utility value for each party. For example, if an agent receives four bars of iron, two bars of silver and two bars of copper, the utility value will be calculated as: (4*4) + (2*3) + (2*2) = 26. User receives the remaining bars (two bars of silver, two bars of copper and four bars of gold) which gets her (4*4) + (2*2) + (2*3)

[1] In a fully distributive negotiation, stating that the situation is “a fixed-pie” is the truth. Therefore, the fixed-pie lie strategy would not gain negotiators any advantage in such a situation.

= 26 points. If they fail to reach an agreement, each party receives 4 points, which makes this their BATNA (Best Alternative To Negotiated Agreement) value.

Table 2. Preference profile for the considered negotiation task.

	Gold	Silver	Copper	Iron
User	4	3	2	1
Agent	1	3	2	4

4 Methods

4.1 Study 1

We used a 2 (Agent Fairness: Fair vs Exploitative) x 2 (Agent Honesty: Truthful vs Liar) design to examine the effects of the agent's fairness and honesty on the participant's behavior in our designed negotiation task.

Fairness (fair vs. exploitative) specifies the value threshold of the deals that are acceptable to the agent. Fair agents would only accept deals that give the agent 50% of the total joint value, and exploitative agents would only settle for 60% or more of the overall joint value.

Honesty (truthful vs. liar) specifies the type of information the agent will provide to the participant. A truthful agent would provide truthful information about its own preferences and a liar will pretend the same preferences as the opponent (fixed-pie lie).

We hypothesized that

H1: Fair agents should require that users make fewer offers to reach agreement compared to exploitative ones.

H2: Liar agents should require that users make fewer offers to reach agreement compared to truthful ones.

We recruited 223 participants using Amazon's Mechanical Turk (MTurk) service as subjects for our study. Participants were presented with a demographic survey followed by instructions on the negotiation task, their preferences, BATNA, and how to interact with the IAGO agent. Next, they were given 10 minutes to negotiate with the agent and come up with an agreement or walk out of the negotiation. Following the negotiation, they were asked to rate how satisfied they were with the results, as well as rate the agent's level of honesty, fairness and strategy.

To ensure that participants understood their priorities in the negotiation, they also completed a manipulation check. Specifically, they were asked to report which items were the most and least valuable for them. From the original 223 participants, 11 failed the manipulation check. These participants were excluded, leaving 212 for analysis.

4.2 Study 2

In this study, we investigated the effectiveness of more exploitative agents. We used a 2 (Agent Fairness: More Exploitative vs Most Exploitative) x 2 (Agent Honesty: Truthful vs Liar). Honesty was manipulated as in study1. Fairness, however, was operationalized in this study as more exploitative (asked for 69% of the overall joint value) and most exploitative (asked for 76% of the overall joint value).

In addition to testing H2 again, we also extent hypothesis H1 to predict that

H3: The more exploitative agent should require that users make fewer offers to reach agreement compared to most exploitative one.

Study flow was the same as the first study and 114 participants were recruited using Amazon's Mechanical Turk (MTurk) service as subjects for this study. From the original 114 participants, 1 failed the manipulation check. This participant was excluded, leaving 113 for analysis.

4.3 Measures

Two main types of measures were considered for evaluating effectiveness of the agent when negotiating with humans: behavioral measures and self-report measures.

Behavioral Measures. The low value of alternative to no agreement (BATNA) in our negotiation task encourages participants to agree on some deal rather than leaving the negotiation with no agreement. However, these agents in only propose (and ultimately accept) deals that gain them equal or more than a specific threshold value (depending on fairness condition). To the extent that participants find this value acceptable, they will propose fewer counter offers. Therefore, the number of counter offers proposed by the participants indicates their dissatisfaction with the agent's proposed value.

Self-report measures. We used self-report measures to assess satisfaction and perceptions of the agent's level of honesty, fairness and strategy. After the negotiation, participants reported their satisfaction with the outcome using a 7-point scale that ranged from 1 (very dissatisfied) to 7 (very satisfied). Then participants rated the agent on honesty, fairness and strategy using 1-7 scales that ranged from 1 (dishonest, unfair, or not-strategic) to 7 (honest, fair, or strategic, respectively).

5 Results

5.1 Study 1

Out of 212 analyzed negotiations, only 6 failed to reach an agreement and settled for their BATNA value. A 2 (Agent Fairness: More Exploitative vs Most Exploitative) x 2 (Agent Honesty: Truthful vs Liar) ANOVA shows that, while there were significant main effects of agent fairness and honesty on user offers ($F(1, 208) = 6.82$, $p = .01$, $d = .36$ and $F(1, 208) = 3.92$, $p = .049$, $d = .29$, respectively), these were qualified by a significant interaction (F(1, 208) = 6.68, p = .01, d = .36). As depicted in Fig. 2, both main effects were driven by users' reactions to the honest exploitative agent compared to the other conditions.

Likewise, the main effects of agent fairness and honesty on satisfaction were either significant (F(1, 208) = 17.61, p < .001, d = .59) or at least trend-like (F(1, 208) = 2.26, p = .13, d = .21), respectively, but again they were qualified by a significant interaction (F(1, 208) = 14.88, p < .001, d = .55). As depicted in Fig. 3, both main effects were again driven by users being less satisfied with the honest exploitative agent compared to the other conditions.

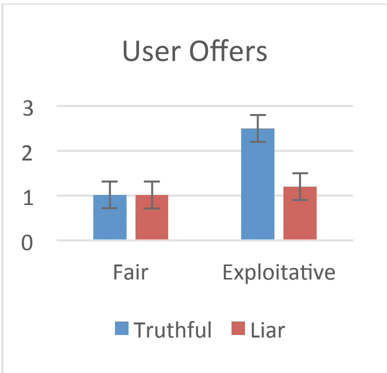

Fig. 2. The number of offers made by users in each condition

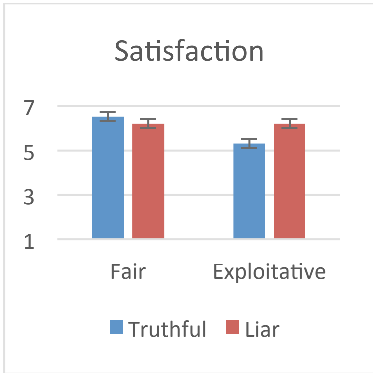

Fig. 3. Self-assessed rating on how satisfied the users were at the end of negotiation

In contrast, users' perceptions of the agent's honesty were unaffected by actual honesty (F(1, 208) = 1.55, p = .22, d = .17) or fairness (F(1, 208) = 0.86, p = .36, d = .13). However, there was a significant interaction (F(1, 208) = 19.61, p < .001, d = .63); as depicted in Fig. 4, the liar fair agent and the honest exploitative agent were both perceived as less honest than the other conditions. Perhaps either appearing to give too much or too little, respectively, can reduce perceived honesty.

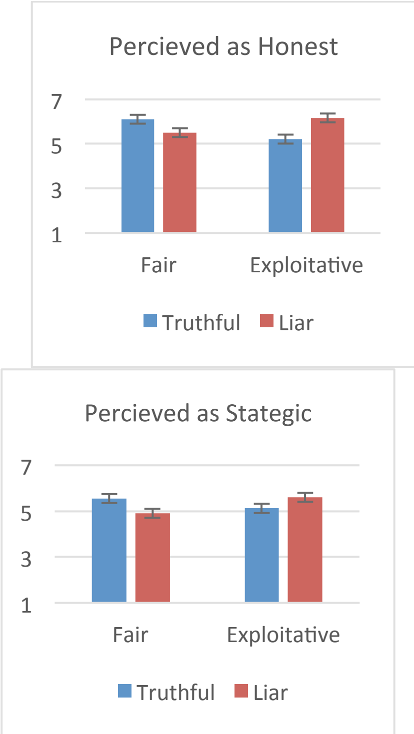

Fig. 4. How honest (left) and strategic (right) agents were perceived by the user (Study 1)

Similarly, users' perceptions of the agent's strategy were unaffected by honesty (F(1, 208) = 0.39, *p* = .53, d = .09) or fairness (F(1, 208) = 0.15, *p* = .70, d = .06), but there was a significant interaction (F(1, 208) = 7.37, *p* = .007, d = .40). As depicted in Fig. 4, the liar fair agent and the honest exploitative agent were also both perceived as less strategic than the other conditions, again perhaps because users thought that either giving too much or too little was not particularly strategic.

5.2 Study 2

Out of 109 analyzed negotiations, 18 failed to reach an agreement and settled for their BATNA value. A 2 (Agent Fairness: More Exploitative vs Most Exploitative) x 2 (Agent Honesty: Truthful vs Liar) ANOVA revealed a significant effect of agent honesty on user offers (F(1, 109) = 7.89, p = .006, d = .54), such that users made fewer offers to agents that lied (M = 1.90, SD = 1.90) than those that told the truth (M = 3.22, SD = 2.94). The main effect of agent fairness and the interaction effect both failed to approach significance (Fs < .14, ps > .71, ds < .06).

However, in this study, participants were no more likely to report feeling satisfied based on the honesty of the agent (F(1, 109) = 0.55, p = .46, d = .14). Instead, there was a trend for participants to report greater satisfaction with the less exploitative of the two agents in this study (M = 4.41, SD = 2.12 vs M = 3.77, SD = 2.00; F(1, 109) =

2.73, p = .10, d = .31). The interaction effect did not approach significance (F(1, 109) = 0.33, p = .57, d = .11).

Likewise, users' perceptions of the agent's honesty were unaffected by actual honesty of the agent (F(1, 109) = 1.05, p = .31, d = .19). Instead, participants perceived the less exploitative of the two agents in this study as marginally more honest (M = 4.67, SD = 1.86 vs M = 4.08, SD = 1.71; F(1, 109) = 3.08, p = .08, d = .33). Again, the interaction effect did not approach significance (F(1, 109) = 0.85, p = .36, d = .18).

In contrast, while there were also trends for main effects of agent fairness and honesty on perceptions of how strategic the agent was (F(1, 109) = 2.03, p = .16, d = .27 and F(1, 109) = 2.56, p = .11, d = .31, respectively), these were qualified by a significant interaction (F(1, 109) = 7.83, p = .006, d = .54). As depicted in Figure 5, both main effects were driven by users perceiving the truthful agent that is relatively less exploitative as more strategic than the other conditions.

6 Discussion and Conclusion

The misrepresentation game is a game theoretic analyses of how a deceptive agent should lie best to gain advantage in a negotiation. Gratch et al. [9] proposed a solution to this problem that the deceiver should feign similar preferences to the opponent for best outcome (fixed-pie lie). This strategy is theoretically proven to be optimal, however, the feasibility of utilizing it by agents in a realistic environment was never tested. Here, we used a publicly available negotiation platform, IAGO, to test this strategy in realistic negotiations with humans.

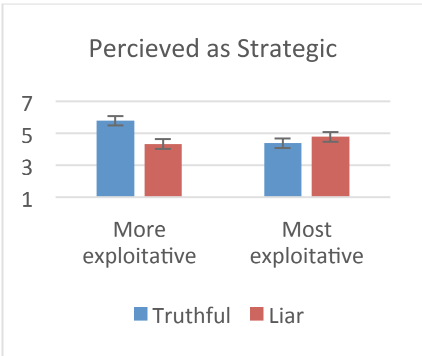

Fig. 5. How strategic agents were perceived by the users (Study 2)

In the first study, we examined the effect of honesty and fairness on user's behavior. We hypothesized that fair agents would require users to make fewer offers to reach agreement compared to the exploitative agent (H1). This hypothesis was confirmed. However, the effect was specifically driven by the truthful exploitative agent, with which users had the hardest time reaching agreement.

Our next hypothesis was that liar agents would require users to make fewer offers to reach agreement compared to truthful agents (H2). Although this hypothesis was confirmed in exploitative condition, participants treated the liar and truthful agents in the fair condition similarly. In exploitative condition, participants also found the liar agent more honest and fairer than the truthful agent, suggesting that there might be a "sweet spot" for deceptive agents. That is, the liar exploitative agent gained the most value, while appearing fair and honest. This result confirms the effectiveness of "fixed-pie lie," particularly in a partially integrative task. In fair condition, however, we found an interesting effect. Even though the liar agent appears to give up a larger share of the overall value, there was a trend for users to be less satisfied with the outcome and find the agent less strategic and less honest. People do not seem to trust an agent who gives up too much or too little, emphasizing the importance of appearing fair (but not too fair) when using "fixed-pie lie" strategy.

Study 1 confirmed the effectiveness of "fixed-pie lie" in partially integrative situations. However, because users showed high levels of satisfaction across the board, we conducted a second study that pushed the boundaries of fairness by using more exploitative agents. We expected the more exploitative agents to require fewer offers from users to reach agreement compared to the most exploitive agents (H3). However, we did not find such effect. It is possible that pushing the boundaries of fairness in this way led to general dissatisfaction. However, Study 2 replicated the finding that liar agents require users to make fewer offers to reach agreement compared to the truthful agent.

Although agents in Study 2 achieved higher gain (only 18 participants walked out of the negotiation out of 109, even though the agents were selfishly asking for 69% or 76% of the overall value), this is probably unrealistic for how agent would do in practice. Since in real-world negotiations users might refuse to negotiate with such agents after this initial exchange, because they were generally dissatisfied, these gains would likely be short-term. In contrast, the liar agent who appears to be fair makes people feel good while gaining more value. The smaller sample size in Study 2 contributed to the effects not reaching traditional levels of significance. For example, three of the trend-like or marginally significant effects would have reached significance if Study 2 had been powered as well as Study 1 (higher satisfaction with the less exploitative agents, higher ratings of honesty for the less exploitative agents, and higher ratings of strategic for the honest agents). Likewise, one trend would have would have been on edge of significance if Study 2 had the sample size of Study 1 (higher ratings of strategic for the less exploitative agents).

These findings emphasize there are clear incentives to creating deceptive agents. An obvious next question is how to avoid or defend against the proliferation of such methods. One often discussed approach is to create ethical guidelines to prevent their creation, yet this is not so simple as it may appear. For example, the type of deception explored in this paper is not actually considered lying by the ethical guidelines that govern human negotiations (see discussion at the end of [9]). Further, the value a negotiator assigns to specific issues is often subjective and difficult to objectively verify. As a result, some negotiation ethicists argue that ethical proscriptions will certainly fail and the only realistic approach is to give people the skills required to detect and defend against such techniques [14]. Indeed, one approach to defeat the fixed-pie lie is to give the opponent what they claim they want to force them to reveal the deception or take a loss to save face. Unfortunately, few of our human participants employed this tactic, and suggests a strong opportunity for using this technology for teaching better skills.

Acknowledgements
This work was supported by the National Science Foundation under grants IIS-1211064 and BCS-1419621, and the U.S. Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

1. T. Baarslag, M. Kaisers, E. H. Gerding, C. M. Jonker, and J. Gratch, "When will negotiation agents be able to represent us The challenges and opportunities for autonomous negotiators," in *Proceedings of the Twenty-sixth International Joint Conference on Artificial Intelligence*, 2017.
2. 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, pp. 345–352.
3. M. Core, D. Traum, H. C. Lane, W. Swartout, J. Gratch, M. van Lent, and S. Marsella, "Teaching negotiation skills through practice and reflection with virtual humans," *Simulation*, vol. 82, no. 11, pp. 685–701, 2006.
4. J. Broekens, M. Harbers, W.-P. Brinkman, C. M. Jonker, K. den Bosch, and J.-J. Meyer, "Virtual reality negotiation training increases negotiation knowledge and skill," in *International Conference on Intelligent Virtual Agents*, 2012, pp. 218–230.
5. J. Gratch, D. DeVault, G. M. Lucas, and S. Marsella, "Negotiation as a challenge problem for virtual humans," in *International Conference on Intelligent Virtual Agents*, 2015, pp. 201–215.
6. M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.
7. N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, M. J. Wooldridge, and C. Sierra, "Automated negotiation: prospects, methods and challenges," *Gr. Decis. Negot.*, vol. 10, no. 2, pp. 199–215, 2001.
8. L. L. Thompson, "Information exchange in negotiation," *J. Exp. Soc. Psychol.*, vol. 27, no. 2, pp. 161–179, 1991.
9. J. Gratch, Z. Nazari, and E. Johnson, "The Misrepresentation Game: How to win at negotiation while seeming like a nice guy," in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, 2016, pp. 728–737.
10. J. Mell and J. Gratch, "IAGO: Interactive Arbitration Guide Online," in

Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, 2016, pp. 1510–1512.
11. Z. Nazari, G. Lucas, and J. Gratch, “Opponent Modeling for Virtual Human Negotiators,” in *Intelligent Virtual Agents*, 2015.
12. T. Baarslag, M. Hendrikx, K. Hindriks, and C. Jonker, “Predicting the performance of opponent models in automated negotiation,” *Proc. - 2013 IEEE/WIC/ACM Int. Conf. Intell. Agent Technol. IAT 2013*, vol. 2, pp. 59–66, 2013.
13. J. Gratch and T. Nguyen, “Misrepresentation Negotiation Games.” , Univ. South. Calif. Inst. Creat. Technol. Play. Vista, CA, 2016.
14. P. R. Reilly, “Was Machiavelli right? Lying in negotiation and the art of defensive self-help,” 2008.