# A study of computational and human strategies in revelation games

**Noam Peled · Ya'akov (Kobi) Gal · Sarit Kraus**

**Abstract** Many negotiations in the real world are characterized by incomplete information, and participants' success depends on their ability to reveal information in a way that facilitates agreements without compromising their individual gain. This paper presents an agent-design that is able to negotiate proficiently with people in settings in which agents can choose to truthfully reveal their private information before engaging in multiple rounds of negotiation. Such settings are analogous to real-world situations in which people need to decide whether to disclose information such as when negotiating over health plans and business transactions. The agent combined a decision-theoretic approach with traditional machine-learning techniques to reason about the social factors that affect the players' revelation decisions on people's negotiation behavior. It was shown to outperform people as well as agents playing the equilibrium strategy of the game in empirical studies spanning hundreds of subjects. It was also more likely to reach agreement than people or agents playing equilibrium strategies. In addition, it had a positive effect on people's play, allowing them to reach significantly better performance when compared to people's play with other people. These results are shown to generalize for two different settings that varied how players depend on each other in the negotiation.

**Keywords** Human–computer negotiation · Opponent modeling · Empirical studies

N. Peled
Gonda Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

Y. Gal (✉)
Department of Information Systems Engineering, Ben-Gurion University, Beer Sheva, Israel
e-mail: gal@eecs.harvard.edu

S. Kraus
Department of Computer Science, Bar-Ilan University, Ramat Gan, Israel

S. Kraus
Institute of Advanced Computer Studies, University of Maryland, College Park, MD, USA

 Springer

## 1 Introduction

In negotiations under incomplete information, people commonly need to make strategic decisions about whether and how to reveal information to others. For example, consider a scenario in which a bank is offering to purchase a struggling company in return for potential job cuts. The unions may not allow the company to accept the offer because they refuse to agree to layoffs. However, if the bank discloses that it is committed to keeping the company afloat, the unions may agree to the buy-out despite the layoffs. On the other hand, revealing goals is often associated with a cost or can be exploited by the other party. In our example, if the bank declares it is committed not to liquidate the company, the unions may demand no layoffs.

The focus of this paper is on negotiation settings in which participants lack information about each other's preferences, often hindering their ability to reach beneficial agreements. Specifically, we study a particular class of such settings we call "revelation games", in which players are given the choice to truthfully reveal private information before commencing in a finite sequence of alternating negotiation rounds. Revealing this information narrows the search space of possible agreements and may lead to agreement more quickly, but may also lead players to be exploited by others.

Revelation games combine two types of interactions that have been studied in the past in the economics literature: Signalling games [25], in which players choose whether to convey private information to each other, and bargaining [19], in which players engage in multiple negotiation rounds. They are analogous to real-world scenarios in which parties may choose to truthfully reveal information before negotiation ensues, such as the example presented above.

Constructing effective agent strategies for such settings is challenging. On the one hand, behavioral economics work has shown that people often follow equilibrium strategies [4] when deciding whether to reveal private information to others. On the other hand, people's bargaining behavior does not adhere to equilibrium [8,18], and computers cannot use such strategies to negotiate well with people [17].

The main contribution of the paper is an agent-design that incorporates information revelation decisions into its negotiation strategy using decision theory and machine learning. The agent tries to determine which proposals people are likely to accept based on past data, or whether they are more or less likely to accept an offer when information is revealed. It combines a prediction model of people's behavior in the game with a decision-theoretic approach to make optimal decisions. The model includes the social factors that affect people's decisions whether to reveal private information, (e.g., the generosity, competitiveness and the selfishness of the offers they make, as measured in the scoring function of the game). In addition, the model included the effects of people's revelation decisions on their negotiation behavior. It evaluated the agent-design with agents playing equilibrium strategies, as well as other people, in two types of revelation games that varied how players depend on each other in the game.

The results showed that the agent was able to outperform human players playing other people, as well as the equilibrium agent. It learned to make offers that were significantly more beneficial to people than the offers made by other people, while not compromising its own benefit, and was able to reach agreement significantly more often than did people as well as the equilibrium agent. In particular, it was able to exploit people's tendency to agree to offers that are beneficial to the agent if people revealed information at the onset of the negotiation. It also positively affected people's play, in that their overall performance was significantly better when paying the agent than when playing other people. The paper thus

has significance for agent designers, showing that (1) people do not adhere to equilibrium strategies when revealing information in negotiation; (2) reasoning about the social factors that affect people's decisions can significantly improve agents' performance as compared to using equilibrium strategies.

This paper extends an initial study reporting on revelation games [22] in several ways. First, it provides a formal equilibrium analysis of revelation games. Second, it analyzes people's behavior in these games, including the extent to which they respond to proposals made by the agent, and how their performance is affected by playing with the agent. Lastly, the empirical analysis includes additional games to the ones originally reported using subjects from different countries. This provides further empirical support to the generalizability of the agent to different demographic groups.

## 2 Related work

Our work is related to a growing line of work in multi-agent systems that use opponent modeling to build agents for repeated negotiation in heterogeneous human-computer settings. These include the KBAgent that made offers with multiple attributes in settings which supported opting out options and partial agreements [20]. This agent used a social utility function to consider the trade-offs between its own benefit from an offer and the probability that it is accepted by people. It used density estimation to model people's behavior following a method suggested by Coehoorn and Jennings for modeling computational agents [6] and approximated people's reasoning by assuming that people would accept offers from computers that are similar to offers they make to each other. Other works employed Bayesian techniques [7,15] or approximation heuristics [16] to estimate people's preferences in negotiation and integrated this model with a pre-defined negotiation or concession strategy to make offers. Bench-Capon [5] provide an argumentation based mechanism for explaining human behavior in the ultimatum game. Work by Rosenfeld and Kraus [23] used Aspiration Adaptation Theory was more useful that other bounded and strictly rational models in quantifying peoples' negotiation preferences. There is also prior work dealing with selective information disclosure in human-agent settings [2,14,24].

None of these works allowed for agents to reveal private information during negotiation. In addition, we are the first to develop a strategic model of people's negotiation behavior that reasons about information revelation, to formalize an optimal decision-making paradigm for agents using this model.

Gal and Pfeffer [11] proposed a model of human reciprocity in a setting consisting of multiple one-shot take-it-or-leave-it games, but did not evaluate a computer agent or show how the model can be used to make decisions in the game. Other works [9,10,13] combined a decision-theoretic approach with a set of rules to adapt to people's negotiation in settings with non-binding agreements. Our work augments these studies in allowing players to reveal private information and in explicitly modeling the effect of revelation on people's negotiation behavior.

Our work is also related to computational models of argumentation, in that people's revelation decisions provide an explanation of the type of offers they make during negotiation. Work in interest-based negotiation has studied different protocols that allows players to reveal their goals in negotiation in a controlled fashion [21,26]. These works assume that agents follow pre-defined strategies for revealing information and do not consider or model human participants.

Lastly, revelation games, which incorporate both signaling and bargaining, were inspired by canonical studies showing that people learn to play equilibrium strategies when they need to signal their private information to others [4]. On the other hand, people's bargaining behavior does not adhere to equilibrium [8,18], and computers cannot use such strategies to negotiate well with people [17]. Our work shows that integrating opponent modeling and density estimation techniques is an effective approach for creating agents that can outperform people as well equilibrium strategies in revelation games.

## 3 Implementation: colored trails

We based our empirical work on a test-bed called Colored Trails [12], which we adapted to model revelation games with 2 rounds, the minimal number that allows an offer to be made by both players. Our revelation game is played on a board of colored squares. Each player has a square on the board that is designated as its goal. The object of the game is to reach the goal square. To move to an adjacent square required surrendering a chip in the color of that square. Players had full view of the board and each others' chips. Both players were shown two possible locations for their goals with associated belief probabilities, but each player could only see its own goal.

We define the no-negotiation alternative score for a player as the score obtained when agreement is not reached. The study included two CT board games. The first board game, called *symmetric*, is shown in Fig. 1a. The initial chip allocation for each player is specified in the Player Chip Display panel. Here, the "me" and "O" icons represent two players, Bob and Alice, respectively. Each player has two possible goals. For example, Bob has two possible goals. Bob's true goal is located four steps to the left of the "me" icon (appearing as a white G square), while Bob's other goal is located four steps below the "me" icon (appearing as



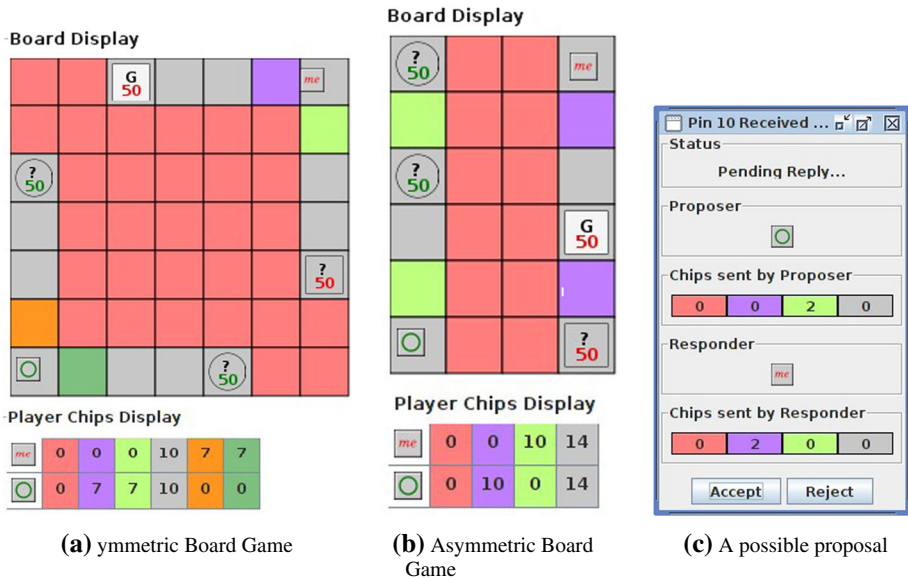**(a)** ymmetric Board Game    **(b)** Asymmetric Board Game    **(c)** A possible proposal

**Fig. 1** Two CT revelation games (shown from Bob's point of view). **a** Symmetric Board Game, **b** asymmetric Board Game, **c** a possible proposal

a grey square outlined with a "?" symbol). In turn, Alice's possible goals are presented as two grey circles outlined with "?" symbols. The board is presented from Bob's point of view. Bob can see its true goal location but Alice does not observe it. Unless Bob chooses to reveal its goal, Alice does not know whether Bob needs a purple or light-green chip to reach the goal. Similarly, Bob cannot observe Alice's true goal location. The number "50" on each goal square represent a 50 % probability that the true goal lies in that square. In the symmetric board, the length of the path between the Alice's location and each of her possible goal squares was equal. Consequently, for the symmetric board, the no-negotiation alternative score of each player does not depend on its true goal location. The analysis from Bob's point of view is similar. In contrast, in the second board, called *asymmetric*, shown in Fig. 1b, one of the players' goal locations is closer to its starting position than the other.

Our CT game progresses in three phases with associated time limits. In the revelation phase (round 0), both players can choose to truthfully reveal their goal to the other player.[1] In the proposal phase (round 1), one of the players is randomly assigned the role of proposer and can offer to exchange a (possibly empty) subset of its chips with a (possibly empty) subset of the chips of the other player. If the responder accepts the offer, the chips are transferred automatically according to the agreement, both participants will automatically be moved as close as possible to the goal square given their chips and the game will end. If the responder rejects (or no offer was received), it will be able to make a counter-proposal (round 2). If the proposal is accepted, the game will end with the agreement result as above. Otherwise, the game will end with no agreement.

At the end of the game, the score for each player is computed as follows: 100 points bonus for reaching the goal; 5 points for each chip left in a player's possession, and 10 points deducted for any square in the path between the players' final position and the goal-square.[2]

The CT game configuration described above is analogous to task settings involving information revelation and incomplete information. We illustrate this analogy using the union agreement scenario presented in the Sect. 1 Goals on the board represent private information that is available to the negotiation parties, such as whether the bank is intending to liquidate the company. Different squares on the board represent different types of sub-tasks, such as agreeing on the number of layoffs, hiring new personnel, and deciding on compensation. Chips correspond to agent capabilities and skills required to fulfill sub-tasks. Traversing a path through the board to the goal square represents reaching agreement that is composed of the various sub-tasks on the paths. The game is interesting because players need to reason about the tradeoff between revealing their goals and providing information to the other player, or not to reveal their goals to possibly receive or ask for more chips than they need. In addition, there is an advantage to the proposer player in the second round, in that it makes the final offer in the game. But the identity of the second proposer is not known at the time that players decide whether to reveal their goals.

## 4 Equilibrium analysis

In this section we formalize the colored trails game described earlier as a repeated game of imperfect information. Each player has a type $t_i$ that represents the true position of its goal on the board.[3] Let $\omega^n$ represent an offer made by a proposer player at round $n \in \{1, 2\}$ in

---

[1] This decision is performed simultaneously by all players, and goals are only revealed at the end of the phase.

[2] This path is computed by the Manhattan distance.

[3] Revealing goals in the game thus corresponds to making players' types common knowledge.

a game. Let $r^n \in \{\text{accept}, \text{reject}\}$ represent the response to $\omega^n$ by a responder player. Let $s_i$ represent the score in the game as described in the previous section. The no-negotiation alternative (NNA) score to player $i$ of type $t_i$ is the score for $i$ in the game given that no agreement was reached. We denote the score for this event as $s_i(\emptyset \mid t_i)$. If no agreement was reached in round 1 of the game, players' final score depends on whether the counter-proposal in round 2 is accepted. If no agreement was reached in round 2 (the last round) of the game, players' NNA score is also their final score in the game. We denote the benefit to player $i$ from $\omega^n$ given that $r^n = accept$ as $\pi_i(\omega^n \mid t_i)$. This is defined as the difference in score to $i$ between an offer $\omega^n$ and the NNA score:

$$\pi_i(\omega^n \mid t_i) = s_i(\omega^n \mid t_i) - s_i(\emptyset \mid t_i) \tag{1}$$

Let $T_i$ denote a set of types for a general player $i$. Let $\text{rev}_i$ denote the event in which $i$ reveals its type $t_i \in T_i$, and let $\text{noRev}_i$ denote the event in which $i$ does not reveal its type. Let $h^n$ denote a history of moves, including for both players $i$ and $j$ their revelation decision at the onset of the game, and the proposals and responses for rounds 1 through $n$. For the remainder of this section we will assume (without loss of generality) that player $j$ is the proposer in round 1, and player $i$ as the proposer in round 2. We denote by $p(t_j \mid h^n)$ the probabilistic belief of player's $i$ on player $j$ being of type $t_j$ (and similarily for player $j$).

Let $a_j^2(\omega, t_j)$ define a strategy for responder $j$ at round 2 of type $t_j$ that accepts any proposal $\omega$ with positive benefit.

$$a_j^2(\omega, t_j) = \begin{cases} \text{accept} & \text{if } \pi_j(\omega \mid t_j) \geq 0 \\ \text{reject} & \text{otherwise} \end{cases} \tag{2}$$

We will abuse notation and write $a_j^2$ when the proposal $\omega$ is clear from context. We extend Eq. 1 to define the benefit to proposer $i$ given the response strategy $a_j^2(\omega, t_j)$.

$$\pi_i(\omega, a_j^2, t_j \mid t_i) = \begin{cases} \pi_i(\omega \mid t_i) & \text{if } a_j^2(\omega, t_j) = \text{accept} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

We denote a proposal from $j$ to $i$ in round 1 as $\omega_{j,i}^1$ and a proposal from $i$ to $j$ in round 2 as $\omega_{i,j}^2$ (also called $i$'s *counter-proposal*). We define several possible strategies for making revelation decisions and proposals in the game, and proceed to show that these strategies are in equilibrium under certain sufficient conditions.

4.1 Revelation by both players

We formalize a revelation equilibrium using backward induction and then provide the conditions under which the equilibrium holds. We assume two possible types for player $j$, denoted $t_{j_1}$ and $t_{j_2}$, and two possible types for player $i$, denoted $t_{i_1}$ and $t_{i_2}$. We begin with the following definitions. Let $\omega_{i,j}^{2*}(t_j)$ denote a proposal in round 2 that maximizes the benefit of player $i$ given that player $j$ is of type $t_j$, the history $h^1 = \{\text{rev}_i, \text{rev}_j, \omega_{j,i}^1, \text{reject}\}$, and that player $j$ uses strategy $a_j^2$ of Eq. 2 to accept proposals.

$$\omega_{i,j}^{2*}(t_j) \in \underset{\omega_{i,j}^2 \in \Omega_{i,j}}{\arg \max} \ \pi_i\left(\omega_{i,j}^2, a_j^2, t_j \mid t_i\right) \tag{4}$$

where (without loss of generality) $\Omega_{i,j}$ is all the set of all the proposals player $i$ can propose to player $j$.

Let $\omega_{j,i}^{1*}(t_i)$ be the proposal that will provide player $j$ with the same benefit as the proposal made by player $i$ in round 2.

$$\omega_{j,i}^{1*}(t_i) \in \arg\max_{\omega_{j,i} \in \Omega_{j,i}} \left( \pi_j(\omega_{j,i} \mid t_j) \text{ s.t. } \pi_i(\omega_{j,i} \mid t_i) = \pi_i(\omega_{i,j}^{2*} \mid t_i) \right) \tag{5}$$

where $\omega_{i,j}^{2*}$ satisfies Eq. 4.

We say that $t_{j_1}$ is *preferable* for $j$ if its benefit from the proposal that $j$ receives in round 2 is greater when it is computed for type $t_{j_1}$ than type $t_{j_2}$, given that $j$ is of type $t_{j_1}$.

$$\pi_j(\omega_{i,j}^{2*}(t_{j_1}) \mid t_{j_1}) \geq \pi_j(\omega_{i,j}^{2*}(t_{j_2}) \mid t_{j_1}) \tag{6}$$

Similarly, we can define when $t_{j_2}$ is *preferable* for $j$. For example, consider the asymmetric game of Fig. 1b. For each player, there are two types in this game, a "weak" type, that is missing two chips to get to the goal, and a "strong" type, that is missing one chip to get to the goal. The weak type will prefer the proposal that assumes it is the weak type (which will provide it with two chips) than the proposal that assumes it is the strong type (which will provide it with one chip). Thus, for each player, the weak type is preferable to the strong type.

We propose the following perfect Bayesian equilibrium.

Round 0: Both players reveal their types for each of their types.

Round 1: We distinguish between two cases. In each case we specify how $i$ updates its beli
In this case, player $i$ updates its beliefs to assign probability 1 to $j$'s type.
In the second case, player $j$ deviates from equilibrium and does not reveal its type. In this case, player $i$ will update its beliefs over $j$ as follows. If type $t_{j_1}$ is preferable to $j$, then player $i$ will update its beliefs that player $j$ is of type $t_{j_2}$.

$$P(t_{j_1} \mid h_1) = 0 \tag{7}$$

(And conversely for the case in which type $t_{j_2}$ is preferable to player $j$). If $i$ revealed its type in round 0 then player $j$ assign probability 1 to $i$'s type. Otherwise, it will arbitrarily assign probability 1 to $t_{i_1}$.
We now describe the actions for players $i$ and $j$ in round 1. Player $j$ makes a proposal $\omega_{j,i}^{1*}(t_i)$ that satisfies Eq. 5. In turn, player $i$ will use the following acceptance strategy $a_i^1(\omega, t_i, t_j)$ such that

$$a_i^1(\omega, t_i, t_j) = \begin{cases} \text{accept} & \text{if } \pi_i(\omega \mid t_i) \geq \pi_i(\omega_{i,j}^{2*}(t_j) \mid t_i) \\ \text{reject} & \text{otherwise} \end{cases} \tag{8}$$

Round 2: Both players do not update their beliefs over each other's actions in round 1.
Player $i$ makes a proposal $\omega_{i,j}^{2*}$ that satisfies Eq. 4. In turn, player $j$ accepts any proposal according to the strategy $a_j^2(\omega, t_j)$ specified in Eq. 2.

## 4.2 Revelation by neither player

In this case neither of the players disclose their true types at the onset of the game. We make the following definitions. Let $E_{h^1}\left[\pi_i\left(\omega_{i,j}^2 \mid t_i\right)\right]$ denote the expected benefit of $i$ for proposal $\omega_{i,j}^2$ given its updated beliefs over $j$'s types and that $j$ accepts beneficial proposals.

$$E_{h^1}\left[\pi_i\left(\omega_{i,j}^2 \mid t_i\right)\right] = \sum_{t_j \in T_j} p\left(t_j \mid h^1\right) \pi_i\left(\omega_{i,j}^2, a_j^2, t_j \mid t_i\right) \tag{9}$$

Let $\omega_{i,j}^{2*}(h^1, t_i)$ be the proposal in round 2 of player $i$ of type $t_i$ that maximizes Eq. 9 given the history $h^1$:

$$\omega_{i,j}^{2*}(h^1, t_i) \in \arg \max_{\omega_{i,j}^2 \in \Omega_{i,j}} E_{h^1} \left[ \pi_i \left( \omega_{i,j}^2 \mid t_i \right) \right] \tag{10}$$

Let $\omega_{j,i}^{1*}$ be the proposal of player $j$ in round 1 that provides it with the same benefit from the counter-proposal made by one of the types of player $i$ in round 2. This proposal does not depend on the type of $j$.

$$\omega_{j,i}^{1*} \in \{\omega_{i,j}^{2*}(h^1, t_{i_1}), \omega_{i,j}^{2*}(h^1, t_{i_2})\} \tag{11}$$

where $\omega_{i,j}^{2*}(h^1, t_i)$ satisfies Eq. 10.

We define a perfect Bayesian equilibrium as follows:

Round 0: Neither player reveals its type for each of its type.

Round 1: We distinguish between two cases. In each case we specify how $i$ updates its beliefs about $j$'s type and which proposal $j$ makes in round 1. In the first case, player $j$ did not reveal its type (i.e., followed the equilibrium strategy). In this case, player $i$ does not update its beliefs over $j$'s types. In the second case, player $j$ revealed its type. In this case, player $i$ updates its beliefs to assign probability 1 to $j$'s type. (And similarly for for player $i$). We now describe players' actions in round 1. Player $j$ will make a proposal $\omega_{j,i}^{1*}$ that satisfies Eq. 11. In turn, player $i$ will use strategy $a_i^1(\omega, t_i)$ such that

$$a_i^1(\omega, t_i) = \begin{cases} \text{accept} & \text{if } \pi_i(\omega \mid t_i) \geq \pi_i(\omega_{i,j}^{2*}(h^1, t_i) \mid t_i) \\ \text{reject} & \text{otherwise} \end{cases} \tag{12}$$

where proposal $\omega_{i,j}^{2*}(h^1, t_i)$ satisfies Eq. 10.

Round 2: Player $i$ does not update its beliefs over $j$'s types (and similarly for player $j$). Player $i$ will make a proposal $\omega_{i,j}^{2*}(h^1, t_i)$ that satisfies Eq. 10. player $j$ accepts any beneficial proposal following Eq. 2.

## 4.3 Sufficient condition for equilibria

We now describe sufficient conditions for which the revelation and non-revelation equilibria hold. For simplicity, we do make the assumption that player $j$ has two types $t_{j_1}$ and $t_{j_2}$. Suppose without loss of generality that $j$ is of type $t_{j_1}$. Let $\omega_{i,j_1}^{2*}$ be the proposal made by player $i$ to player $j$ of type $t_{j_1}$ that satisfies Eq. 4 (when player $j$ reveals its type) and let $\omega_{i,j}^{2*}(h^1)$ be the proposal that satisfies Eq. 10 (when player $j$ does not reveal its type) and that $h^1$ is the history up until round 2. We will postulate two propositions:

**Proposition 1** *If one of $j$'s type is preferable (Eq. 6) for $j$, then the strategies specified in Sect. 4.1 constitute a perfect Bayesian equilibrium in which both players $i$ and $j$ reveal their types for each type in round 0.*

**Proposition 2** *If the below conditions hold, then the strategies specified in Sect. 4.2 constitute a perfect Bayesian equilibrium in which both players do not reveal their types for each of their types in round 0.*

*The first condition says $j$ prefers not to disclose its type. Formally, for every type $t_j$ and $t_i$, we say that*

$$\pi_j(\omega_{i,j}^{2*}(t_j) \mid t_j) \le \pi_j\left(\omega_{i,j}^{2*}(h^1, t_i) \mid t_j\right) \tag{13}$$

*The second condition says that any proposal that is greater or equal to i in round 1 is (at least) worse off for $j$ for each of its types. Formally, we say that for every type $t_j$ and $t_i$, the following holds.*

$$\forall \omega_{j,i}^1 \text{ if } \pi_i(\omega_{j,i}^1 \mid t_i) \ge \pi_i\left(\omega_{i,j}^{2*}(h^1, t_i) \mid t_j\right) \text{ then } \pi_j(\omega_{j,i}^1 \mid t_j) \le \pi_j\left(\omega_{i,j}^{2*}(h^1, t_i) \mid t_j\right) \tag{14}$$

In the Appendix, we provide a proof to the Proposition above, and also show that the sufficient conditions hold in the board games in our study.

### 4.4 Adaptation of equilibria to CT game

We will exemplify the equilibrium strategies in the revelation games for the two boards in Fig. 1a, b and demonstrate that they are in equilibrium. In this analysis, each of the two types of each player is assigned a 50 % prior probability. For expository convenience we assume that the "me" player is the first proposer in the game, as shown in the figure.[4]

We begin with the symmetric game in Fig. 1a. Here, both players are located at equal distance of each of their goals. The no-negotiation alternative for both players is 80 points. The "me" player has 24 chips at the onset of the game, and is missing one purple chip to reach the goal square from its initial location. In this game we describe an equilibrium in which neither player reveals its goal.

Round 0: Neither of the players reveals its goal.

Round 1: The "me" player will propose to give 21 chips (7 olive-green, 7 orange chips and 7 grey chips) in return for two chips (one purple and one green chip). This proposal will yield a score of 105 for the "me" player and a score of 295 for the "O" player. The "O" player accepts any proposal that provides it with higher benefit than the proposal it will make in round 2.

Round 2: The "O" player makes the same offer as in round 1. The "me" player accepts any proposal that provides it with positive benefit.

In contrast to the symmetric game, in the asymmetric game the distance of each player from the goal depends on its true goal location, thus the no-negotiation alternative score depends on the type of the player. We therefore distinguish between two possible types of players, *weak* and *strong*. The weak player is missing two chips to get to the goal, and its no-negotiation alternative score is 70 points. The strong player is missing a single chip to get to the goal, and its no-negotiation alternative score is 90 points. In Fig. 1b, which shows players' starting positions on the game board, the "me" player's type is strong, because its starting position is located closer to its true goal location (the highlighted square marked "G") than to its other possible goal location. For the asymmetric game of Fig. 1b we present an equilibrium strategy that depend on the type of the "me" player.

---

[4] In practice, the identity of the first proposer was determined stochastically according to a known probability distribution. In practice this did not affect the equilibrium strategies that are described in this section.

Round 0: Both of the player disclose their goals.

Round 1: The proposal made by the "me" player depends on its known type. A weak player will propose 10 green chips and 11 gray chips to player "O" in return for 2 purple chips. A strong player will propose 10 green chips and 12 gray chips to player "O" in return for 1 purple chip. The "O" player accepts any proposal that provides it with higher benefit than the proposal it will make in round 2.

Round 2: The "O" player proposes the offer in round 1 for the the "me" player, which depends on the "me" player's type. The "me" player will accept any proposal with positive benefit.

## 5 The SIGAL agent

In this section we describe the Sigmoid Acceptance Learning Agent (SIGAL). SIGAL uses a decision-theoretic approach to negotiate in revelation games, that is based on a model of how humans make decisions in the game. Before describing the strategy used by SIGAL we make the following definitions.

For the remainder of this section, we assume that the SIGAL agent (denoted $a$) is playing a person (denoted $p$). Let $\omega_{a,p}^n$ represent an offer made by the agent to the person in round $n$ and let $r_p^n$ represent the response of the person to $\omega_{a,p}^n$. The expected benefit to SIGAL from $\omega_{a,p}^n$ given history $h^{n-1}$ and SIGAL's type $t_p$ is denoted $E_a\left(\omega_{a,p}^n \mid h^{n-1}, t_a\right)$. Let $p(r_p^n = \text{accept} \mid \omega_{a,p}^n, h^{n-1})$ denote the probability that $\omega_{a,p}^n$ is accepted by the person given history $h^{n-1}$.

We now specify the strategy of SIGAL for the revelation game defined in Sect. 3. The strategy assumes there exists a model of how humans make and accept offers in both rounds. We describe how to estimate the parameters of this model in Sect. 5. We begin by describing the negotiation strategies of SIGAL for rounds 2 and 1.

**Round 2:** If SIGAL is the second proposer, its expected benefit from an offer ($\omega_{a,p}^2$) depends on its model of how people accept offers in round 2, encapsulated in the probability $p(r_p^2 = \text{accept} \mid \omega_{a,p}^2, h^1)$. The benefit to SIGAL is

$$
\begin{aligned}
E_a &\left(\omega_{a,p}^2 \mid h^1, t_a\right) \\
&= \pi_a(\omega_{a,p}^2 \mid t_a) \cdot p(r_p^2 = \text{accept} \mid \omega_{a,p}^2, h^1) \\
&\quad + \pi_a(\emptyset \mid t_a) \cdot p(r_p^2 = \text{reject} \mid \omega_{a,p}^2, h^1)
\end{aligned}
$$

Here, the term $\pi_a(\emptyset \mid t_a)$ represents the benefit to SIGAL from the NNA score, which is zero. SIGAL will propose an offer that maximizes its expected benefit in round 2 out of all possible proposals for this round.

$$
\omega_{a,p}^{2*} = \arg \max_{\omega_{a,p}^2} E_a\left(\omega_{a,p}^2 \mid h^1, t_a\right) \tag{15}
$$

If SIGAL is the second responder, its optimal action is to accept any proposal from the person that gives it positive benefit as described in the equilibrium strategy (Eq. 2). Let $r_a^{2*}(\omega_{p,a}^2 \mid h^1)$ denote the response of SIGAL to offer $\omega_{p,a}^2$, defined as

$$
r_a^{2*}(\omega_{p,a}^2 \mid h^1) = \begin{cases} \text{accept} & \pi_a(\omega_{p,a}^2 \mid t_a) > 0 \\ \text{reject} & \text{otherwise} \end{cases} \tag{16}
$$

where $\pi_a(\omega_{p,a}^2 \mid t_a)$ is defined in Eq. 1. The benefit to SIGAL from this response is defined as

$$
\pi_a\left(r_a^{2*} \mid \omega_{p,a}^2, h^1, t_a\right)
$$
$$
= \begin{cases} \pi_a(\omega_{p,a}^2 \mid t_a) & r_a^{2*}(\omega_{p,a}^2 \mid h^1) = \text{accept} \\ \pi_a(\emptyset \mid t_a) & \text{otherwise} \end{cases} \tag{17}
$$

**Round 1:** If SIGAL is the first proposer, its expected benefit from making a proposal $\omega_{a,p}^1$ depends on its model of the person: If the person accepts $\omega_{a,p}^1$, then the benefit to SIGAL is just $\pi_a(\omega_{a,p}^1 \mid t_a)$. If $(\omega_{a,p}^1)$ is rejected by the person, then the benefit to SIGAL depends on the counter-proposal $\omega_{p,a}^2$ made by the person in round 2, which itself depends on SIGAL's model $p(\omega_{p,a}^2 \mid h^1)$ of how people make counter-proposals. The expected benefit to SIGAL from behaving optimally as a second responder for a given offer $\omega_{p,a}^2$ is denoted $E_a(resp^2 \mid h^1, t_a)$, and defined as

$$
E_a(resp^2 \mid h^1, t_a)
$$
$$
= \sum_{\omega_{p,a}^2} p(\omega_{p,a}^2 \mid h^1) \cdot \pi_a(r_a^{2*} \mid \omega_{p,a}^2, h^1, t_a) \tag{18}
$$

where $\pi_a(r_a^{2*} \mid \omega_{p,a}^2, h^1, t_a)$ is defined in Eq. 17.

Its expected benefit from $\omega_{a,p}^1$ is:

$$
E_a\left(\omega_{a,p}^1 \mid h^0, t_a\right)
$$
$$
= \pi_a(\omega_{a,p}^1 \mid t_a) \cdot p(r_p^1 = \text{accept} \mid \omega_{a,p}^1, h^0)
$$
$$
+ E_a(resp^2 \mid h^1, t_a) \cdot p(r_p^1 = \text{reject} \mid \omega_{a,p}^1, h^0) \tag{19}
$$

where $h^1 = \{h^0, \omega_{a,p}^1, r_p^1 = \text{reject}\}$. SIGAL will propose an offer in round 1 that maximizes its expected benefit in this round:

$$
\omega_{a,p}^{1*} = \arg\max_{\omega_{a,p}^1} E_a\left(\omega_{a,p}^1 \mid h^0, t_a\right) \tag{20}
$$

If SIGAL is the first responder, it accepts any offer that provides it with a larger benefit than it would get from making the counter-proposal $\omega_{a,p}^{2*}$ in round 2, given its model of how people respond to offers in round 2:

$$
r_a^{1*}(\omega_{p,a}^1 \mid h^0) = \begin{cases} \text{accept} & \pi_a(\omega_{p,a}^1 \mid t_a) > \\ & E_a\left(\omega_{a,p}^{2*} \mid h^1, t_a\right) \\ \text{reject} & \text{otherwise} \end{cases} \tag{21}
$$

Here, $h^1 = \{h^0, \omega_{p,a}^1, r_a^1 = \text{reject}\}$, $\pi_a(\omega_{p,a}^1 \mid t_a)$ is defined in Eq. 1 and $E_a\left(\omega_{a,p}^{2*} \mid h^1, t_a\right)$ is the benefit to SIGAL from making an optimal proposal $\omega_{a,p}^{2*}$ at round 2, as defined in Eq. 15.

Let $\pi_a\left(r_a^{1*} \mid \omega_{p,a}^1, h^0, t_a\right)$ denote the benefit to SIGAL from its response to offer $\omega_{p,a}^1$ in round 1. If SIGAL accepts this offer, it receives the benefit associated with $\omega_{p,a}^1$. If it rejects

this offer, it will receive the expected benefit $E_a\left(\omega_{a,p}^{2*} \mid h^1, t_a\right)$ from making an optimal counter-proposal at round 2:

$$
\pi_a\left(r_a^{1*} \mid \omega_{p,a}^1, h^0, t_a\right)
$$
$$
= \begin{cases} \pi_a(\omega_{p,a}^1 \mid t_a) & r_a^{1*}(\omega_{p,a}^1 \mid h^0) = \text{accept} \\ E_a\left(\omega_{a,p}^{2*} \mid h^1, t_a\right) & \text{otherwise} \end{cases} \tag{22}
$$

The expected benefit to SIGAL as a responder in round 1 is denoted as $E_a\left(resp^1 \mid h^0, t_a\right)$. This benefit depends on its model of all possible offers made by people for each type, given that SIGAL responds optimally to the offer.

$$
E_a\left(resp^1 \mid h^0, t_a\right) = \sum_{t_p \in T_p} p(t_p \mid h^0) \cdot
$$
$$
\left( \sum_{\omega_{p,a}^1} p(\omega_{p,a}^1 \mid t_p, h^0) \cdot \pi_a(r_a^{1*} \mid \omega_{p,a}^1, h^0, t_a) \right) \tag{23}
$$

Note that when the person reveals his/her type at round 0, this is encapsulated in the history $h^0$, and $p(t_p \mid h^0)$ equals 1 for the person's true type. Otherwise $p(t_p \mid h^0)$ equals the probability $p(t_p)$.

**Round 0:** In the revelation round SIGAL needs to decide whether to reveal its type. Let $E_a(h^0, t_a)$ denote the expected benefit to SIGAL given that $h^0$ includes a revelation decision for both players and that $t_a$ is the type of agent. This benefit depends on the probability that SIGAL is chosen to be a proposer ($p(prop)$) or responder ($p(resp)$) in round 1:

$$
E_a(h^0, t_a) = p(resp) \cdot E_a\left(resp^1 \mid h^0, t_a\right)
$$
$$
+ p(prop) \cdot E_a(\omega_{a,p}^{1*} \mid h^0, t_a) \tag{24}
$$

Here, $\omega_{a,p}^{1*}$ is the optimal proposal for SIGAL in round 1, and $E_a\left(\omega_{a,p}^{1*} \mid h^0, t_a\right)$ is the expected benefit associated with this proposal, defined in Eq. 19.

Because players do not observe each other's revelation decisions, the expected benefit for a revelation decision $\phi_a$ of the SIGAL agent sums over the case where people revealed their type (i.e., $\phi_a = t_p$) or did not reveal their type (i.e., $\phi_a = null$). We denote $p(\phi_a = t_p)$ as the probability that the person revealed its type $t_p$, and $p(\phi_p = null)$ as the probability that the person did not reveal its type $t_p$.

$$
E_a\left(\phi_a\right) = \sum_{t_p \in T_p} \left[ p(\phi_p = t_p) \cdot E_a\left(h^0 = \{\phi_a, \phi_p = t_p\}, t_a\right) \right.
$$
$$
\left. + p(\phi_p = null) \cdot E_a\left(h^0 = \{\phi_a, \phi_p = null\}, t_a\right) \right] \tag{25}
$$

Given that SIGAL is of type $t_a \in T_a$, it reveals its type only if its expected benefit from revelation is greater or equal to not revealing its type:

$$\phi_a^* = \begin{cases} t_a & E_a \left( \phi_a = t_a \right) \geq \\ & E_a \left( \phi_a = null \right) \\ null & \text{otherwise} \end{cases} \tag{26}$$

The value of the game for SIGAL for making the optimal decision whether to reveal its type is defined as $E_a \left( \phi_a^* \right)$.

Lastly, we wished SIGAL to take a risk averse approach to making decisions in the game. Therefore SIGAL used a convex function to represent its utility in the game from an offer $\omega^n$, which modified Eq. 1.

$$\pi_a'(\omega^n \mid t_a) = \frac{\pi_a(\omega^n \mid t_a)^{(1-\rho)}}{1 - \rho} \tag{27}$$

The strategy used by SIGAL is obtained by "plugging in" the risk averse utility $\pi_a'(\omega^n \mid t_a)$ instead of $\pi_i(\omega^n \mid t_i)$.

## 6 Modeling human players

In this section we describe a model of people's behavior used by SIGAL to make optimal decisions in the game. We assume that there is a training set of games played by people, as we show in the next section.

### 6.1 Accepting proposals

We modeled people's acceptance of proposals in revelation games using a stochastic model that depended on a set of features. These comprised past actions in the game (e.g., a responder may be more likely to accept a given offer if it revealed its type as compared to the case in which it did not reveal its type) as well as social factors (e.g., a responder player may be less likely to accept a proposal that offers more benefit to the proposer than to itself).[5]

Let $\omega_{i,j}^n$ represent a proposal from a player $i$ to a player $j$ at a round $n$. We describe the following features that affect the extent to which player $j$ will accept proposal $\omega_{i,j}^n$. These features are presented from the point of view of proposer $i$, therefore we assume that the type of the proposer $t_i$ is known, while the type of the responder $t_j$ is known only if $j$ revealed its type. We first detail the features that relate to players' decisions whether to reveal their types.

- $REV_j^0$. Revelation by $j$. This feature equals 1 if the responder $j$ has revealed its type and 0 otherwise. The superscript 0 indicates this feature is relevant to the revelation phase, which is round 0.
- $REV_i^0$. Revelation by $i$. This feature equals 1 if the proposer has revealed its type $t_i$.

We now describe the set of features relating to social factors of the responder player $j$.

- $BEN_j^n$. Benefit to $j$. The benefit to $j$ from proposal $\omega_{i,j}^n$ in round $n$. This measures the extent to which the proposal $\omega_{i,j}^n$ is generous to the responder. In the case where $j$ revealed its type, this feature equals $\pi_j(\omega_{i,j}^n \mid t_j)$ and computed directly from Eq. 1. Otherwise, the value of this feature is the expected benefit to the responder from $\omega_{i,j}^n$ for all possible responder types $T_j$:

---

$$\sum_{t_j \in T_j} p(t_j \mid h^{n-1}) \cdot \pi_j(\omega_{i,j}^n \mid t_j)$$

- $AI_i^n$. Advantageous inequality of $i$. The difference between the benefit to proposer $i$ and responder $j$ that is associated with proposal $\omega_{i,j}^n$. This measures the extent to which proposer $i$ is competitive, in that $\omega_{i,j}^n$ offers more for $i$ than for $j$. This feature equals the difference between $\pi_i(\omega_{i,j}^n, \text{accept} \mid t_i)$ and $BEN_j^n$.

To capture the way the behavior in round $n = 1$ affects the decisions made by participants in round $n = 2$, we added the following features that refer to past offers.

- $P.BEN_j^n$. Benefit to $j$ in the previous round. This feature equals $BEN_j^1$ if $n = 2$, and 0 otherwise.
- $P.BEN_i^n$. Benefit to proposer $i$ in the previous round. This feature equals $\pi_i(\omega_{i,j}^1, \text{accept} \mid t_i)$ if $n = 2$ and 0 otherwise.

To illustrate, consider the asymmetric CT board game shown in Fig. 1b. Alice is missing two green chips to get to the goal and Bob is missing 1 purple chip to get to the goal. Suppose Bob is the first proposer (player $i$) and that Alice is the first responder (player $j$), and that Bob revealed its goal to Alice, so its type is common knowledge, while Alice did not reveal her goal. We thus have that $REV_j^0 = 0$ and $REV_i^0 = 1$. Alice's no-negotiation alternative (NNA) score, $s_j(\emptyset)$, is 70 points and Bob's NNA score is 90 points.

According to the offer shown in the Figure, Bob offered two green chips to Alice in return for two purple chips. If accepted, this offer would allow Alice to get to the goal in 5 steps, so she will have 19 chips left at the end of the game, worth $19 \cdot 5 = 95$ points. Similarly, Bob will have 21 chips left at the end of the game, worth 105 points. Both will also earn a bonus of 100 points for getting to the goal. Therefore we have that $BEN_j^1 = 95 + 100 - 70 = 125$. Similarly, Bob's benefit from this proposal is $105 + 100 - 90 = 115$ points. The difference between the benefit to Bob and to Alice is $-10$, so we have that $AI_i^1 = -10$. Lastly, because the offer is made in round 1, we have that $P.BEN_j^1 = P.BEN_i^1 = 0$. This offer is more generous to Alice than it is to Bob.

Suppose now that Alice rejects this offer and makes a counter proposal in round 2, that proposes one purple chip to Bob in return for four greens. In this example, Alice is using her knowledge of Bob's type to make the minimal offer that would allow Bob to reach the goal while providing additional benefit to Alice. Alice is the proposer (player $i$) and Bob is the responder (player $j$). Recall that Bob has revealed its goal while Alice did not, so we have $REV_j^0 = 1$ and $REV_i^0 = 0$. Using a similar computation from before, we get that Bob's score from the counter proposal is 190 points. Therefor we have that $BEN_j^2 = 190 - 90 = 100$. Alice's benefit from the counter-proposal is $210 - 70 = 140$, therefore we have that $AI_i^2 = 140 - 100 = 40$. The last features in the example capture the benefit to both players from the proposal made in the first round to Alice and Bob, so we have $P.BEN_j^2 = 125$, and $P.BEN_i^2 = 115$.

### 6.1.1 Social utility function

We model the person as using a social utility function to decide whether to accept proposals in the game. This social utility depends on a weighted average of the features defined above. We define a transition function, $T^n$, that maps an offer $\omega^n$ and history $h^{n-1}$ to an (ordered) set of feature values $x^n$ as follows.[6]

---

[6] These weights are estimated from data using statistical techniques as described in the following section.

$$x^n = \left( REV_j^0, REV_i^0, BEN_j^n, AI_i^n, P.BEN_j^n, P.BEN_i^n \right)$$

To illustrate, in the example above, we have that $x^1 = (0, 1, 125, -10, 0, 0)$ and $x^2 = (1, 0, 100, 40, 125, 115)$.

Let $u(x^n)$ denote the social utility function which is defined as the weighted sum of these features. To capture the fact that a decision might be implemented noisily, we use a sigmoid function to describe the probability that people accept offers, in a similar way to past studies for modeling human behavior [11]. We define the probability of acceptance for a particular features values $x^n$ by a responder to be

$$p(r_i^n = \textsf{accept} \mid \omega^n, h^{n-1}) = \frac{1}{1 + e^{-u(x^n)}} \tag{28}$$

where $x^n = T^n(\omega^n, h^{n-1})$. In particular, the probability of acceptance converges to 1 as $u(x^n)$ becomes large and positive, and to 0 as the utility becomes large and negative. We interpret the utility to be the degree to which one decision is preferred. Thus, the probability of accepting a proposal is higher when the utility is larger.

### 6.1.2 Estimating weights

To predict how people respond to offers in the game, it is needed to estimate the weights in their social utility function in a way that best explains the observed data. In general, we need to model the probability that an offer is accepted for any possible instantiation of the history. The number of possible proposals in round 1 is exponential in the combined chip set of players.[7] It is not possible to use standard density estimation techniques because many such offers were not seen in the training set or were very rare. Therefore, we employed a supervised learning approach that assumed people used a noisy utility function to accept offers that depended on the features defined above. Let $\Omega_{i,p}$ denote a data set of offers proposed by some participant $i$ to a person $p$.[8] For each offer $\omega_{i,p}^n \in \Omega_{i,p}$ let $y(r_p^n \mid \omega_{i,p}^n)$ denote an indicator function that equals 1 if the person accepted proposal $\omega_{i,p}^n$, and zero otherwise. The error of the predictor depends on the difference between $y(r_p^n \mid \omega_{i,p}^n)$ and the predicted response $p(r_p^n = \textsf{accept} \mid \omega_{a,p}^n, h^{n-1})$, as follows:

$$\sum_{\omega_{i,p}^n \in \Omega_{i,p}} \left( p(r_j^n = \textsf{accept} \mid \omega_{i,p}^n, h^{n-1}) - y(r_j^n \mid \omega_{i,p}^n) \right)^2 \tag{29}$$

where $p(r_j^n = \textsf{accept} \mid \omega_{i,j}^n, h^{n-1})$ is defined in Eq. 28.

We used a standard Genetic algorithm to estimate weight values for the features of people's social utility that minimize the aggregate error in the training set. To avoid over-fitting the training set, we used a held-out cross-validation set consisting of 30 % of the data. We chose the instance with minimal error (on the training set) in the generation that corresponded to the smallest error on the cross-validation set. We used ten-fold cross-validation, repeating this process ten times, each time choosing different training and testing sets, producing ten candidate instances. To pick the best instance, we computed the value of the game $E_a(\phi_a^*)$ for SIGAL for each of the learned models, where $\phi_a^*$ is defined in Eq. 26. This is the expected benefit for SIGAL given that it chooses optimal actions using a model of people that corresponds to the feature values in each instance.

---

[7] In one of the boards we studied the number of possible offers that provided the same benefit to both players was about 27,000, out of a total of $2^{24}$ possible offers.

[8] We explain how we collected this data set in the Sect. 7

## 6.2 Proposing and revealing

This section describes our model of how people make proposals in revelation games and reason about whether to reveal information.

### 6.2.1 First proposal model

We used standard density estimation techniques (histograms) to predict people's offers for different types. Based on the assumption that proposals for the first round depend on the proposer's type and its decision whether to reveal, we divided the possible proposals to equivalence classes according to the potential benefit for the proposer player, and counted how many times each class appears in the set. Let $p(\omega^1_{p,j} \mid t_p, \phi_i)$ denote the probability that a human proposer of type $t_p$ offers $\omega^1_{p,j}$ in round 1. Let $N_{t_p, \phi_p}(\pi_p(\omega^1_{p,j} \mid t_p))$ denote the number of proposals in round 1 which gives the human a benefit of $\pi_p(\omega^1_{p,j} \mid t_p)$, given the human is of type $t_p$ and its revelation decision was $\phi_p$. Let $N_{t_p, \phi_p}(\Omega^1_{p,j})$ denote the number of proposals in round 1 in this subset. $p(\omega^1_{p,j} \mid t_p, \phi_p)$ is defined as:

$$p(\omega^1_{p,j} \mid t_p, \phi_p) = \frac{N_{t_p, \phi_p}\left(\pi_p(\omega^1_{p,j} \mid t_p)\right)}{N_{t_p, \phi_p}(\Omega^1_{p,j})} \tag{30}$$

### 6.2.2 Counter-proposal model

According to our model, a player's proposal in the second round also depends on the history, this two dimensional probability density function tends to be too sparse to calculate it directly as described in Sect. 6.2.1. Inspired by studies showing that people engage in tit-for-tat reasoning [27] we used this principal to model the counter-proposals made by people. We assumed that a responder player $i$ will be proposed offer $\omega^2_{p,i}$ by a human player in the second round with benefit $\pi_i(\omega^2_{p,i} \mid t_i)$ that is equal to the benefit $\pi_p(\omega^1_{i,p} \mid t_p)$ from offer $\omega^1_{i,p}$ made to the person in the first round, when the human was a responder. For example, suppose that Bob is the proposer in round 1 and propose to Alice a benefit of 125. According to the model, if Alice rejects the offer she will propose Bob a counter-proposal that provides Bob with the same benefit, 125. Note that this does not assume that the proposal will provide Alice with the same benefit she got from Bob in the proposal from round 1. Formally, let $N_{\Omega^2_{p,i}}(\pi_p(\omega^1_{i,p} \mid t_p))$ denote the number of counter-proposals $\omega^2_{p,i}$ which give benefit $\pi_p(\omega^1_{i,p} \mid t_p)$. We assume that there always exists at least one proposal that meets this criterion, i.e., $N_{\Omega^2_{p,i}}(\pi_p(\omega^1_{i,p} \mid t_p)) \neq 0$. The "tit for tat" heuristic is as follows:

$$p(\omega^2_{p,i} \mid h^1) = \begin{cases} 0 & \pi_i(\omega^2_{p,i}) \neq \pi_p(\omega^1_{i,p}) \\ 1/N_{\Omega^2_{p,i}}(\pi_p(\omega^1_{i,p} \mid t_P)) & \text{otherwise} \end{cases} \tag{31}$$

This heuristic is used in Eq. 18 to facilitate the computation of the expected benefit from SIGAL as a responder in round 1.

Lastly, we detail the model used by SIGAL to predict whether the person reveals its goal. Let $N_{t_p}$ denote the number of instances in which people were of type $t_p$, and let $N_{t_p}(\phi_p)$ denote the number of times that people of type $t_p$ chose to reveal their type. The probability that a human player $p$ revealed its type $t_p$ is defined as:

$$p(\phi_p \mid t_p) = \frac{N_{t_p}(\phi_p)}{N_{t_p}} \tag{32}$$

## 7 Empirical methodology

In this section we describe the methodology we used in order to learn the parameters of the model of how people play revelation games, and to evaluate it. For these purposes we recruited 260 students enrolled in a computer science or software engineering program at several universities and colleges. An additional 143 people were recruited using the Amazon Turk framework. All subjects received an identical tutorial on revelation games that was exemplified on a board (not the boards used in the study). Actual participation was contingent on successfully answering a set of basic comprehension questions about the game. Participants were seated in front of a terminal for the duration of the study, and could not speak to any of the other participants. Each participant played two revelation games on different boards.

The boards in the study fulfilled the following conditions at the onset of the game: (1) There were two goals for each player; (2) Every player lacked one or two chips to reach each of its possible goals; (3) Every player possessed the chips that the other needed to get to each of its possible goals; (4) There existed at least one exchange of chips which allowed both players to reach each of their possible goals; (5) the goals were distributed with a probability of 50 % for both players.

We used the asymmetric and symmetric boards shown in Fig. 1a, b. Participants played both symmetric and asymmetric boards in random order. They engaged in a neutral activity (answering demographic questions) between games to minimize the effects of their behavior in the first game on their behavior in the second game. The participant chosen to be the proposer in the first game was randomly determined, and participants switched roles in the second game, such that the proposer in the first game was designated as the responder in the second game. A central server (randomly) matched each participant with a human or an agent counterpart for each game. The identity of each participant was not disclosed. We collected players' proposals, responses and revelation decisions for all of the games played. To avoid deception all participants were told they would be interacting with a computer or a person. Participants received fixed compensation (course credit) for participating in the experiment.[9]

We divided subjects into four disjoint pools. The first pool consisted of people playing other people (66 games). The second pool consisted of people playing a computer agent that used a randomized strategy to make offers and responses (170 games). The purpose for this pool was to collect people's actions for diverse situations, for example, their response to offers that were never made by other people. Two thirds (44 games) of the data from the first pool and the data from the second pool were used for training a model of people's behavior. The third pool consisted of people playing the SIGAL agent (238 games). The fourth pool (118 games) consisted of people playing an agent using several of the equilibrium strategies defined in Sect. 4 to play revelation games. The equilibrium agent adapted a non-revelation and revelation strategy equilibrium as specifed in Sects. 4.1 and 4.2 for each game.

---

[9] Our goal was to build an agent that negotiates well with people, not to explain people's incentives, therefore fixed compensation was sufficient.

## 8 Results and discussion

The performance of SIGAL was measured by comparing its performance against people (the third pool) with people's play against other people (the first pool). We list the number of observations and means for each result. All results reported in this section are statistically significant in the $p < 0.05$ range using $t$ tests for normally distributed data and two-sample Mann–Whitney rank tests for non-normally distributed data.

8.1 Analysis: general performance

We first present a comparison of the performance of SIGAL and people. Figure 2 shows the average benefit (the difference in score between agreement and the no-negotiation alternative score) for different roles (proposer and responder). As shown by the Figure, the SIGAL agent outperformed people in all roles. In addition, SIGAL was also more successful at reaching agreements than were people. Figure 3 shows the percentage of offers accepted by people for the different roles. As shown by the Figure, proposals made by SIGAL in round 1 were accepted 62 % of the time, while proposals made by people in round 1 were accepted only 49 % of the time. This difference is more pronounced in round 2, in which proposals made by SIGAL were accepted 83 % of the time, while offers made by people in round 2 were only accepted 63 % of the time. If an offer is rejected at this last round, the game ends without agreement. This striking difference shows that SIGAL learned to make good offers at critical points in the game.

As shown in Figure 2 SIGAL also outperformed the equilibrium agent in both rounds. The equilibrium agent was fully strategic and assumed the other player was unboundedly rational. Although not shown in the Figure, it made very selfish offers in the last round, offering only 25 average benefit points to people and 215 benefit points to itself. Most of these offers (54 %) were rejected. In the first round, it made offers that were highly beneficial to people, offering 219 average points to people and 20 to itself. Most of these offers (82 %) were accepted, but the small benefit it incurred in these proposals not aid its performance.
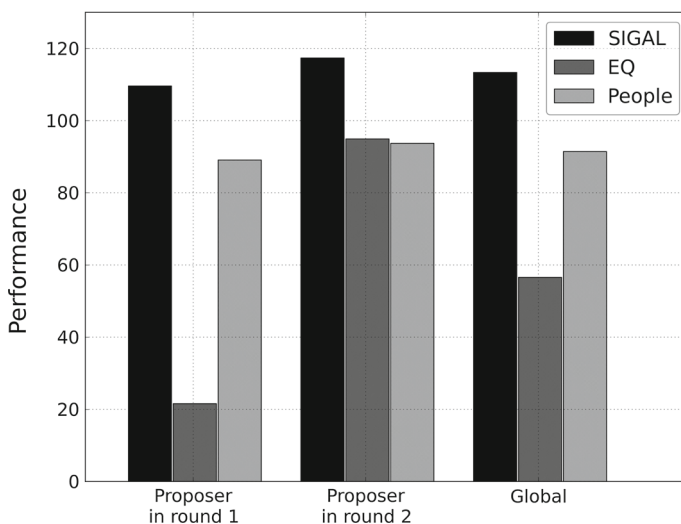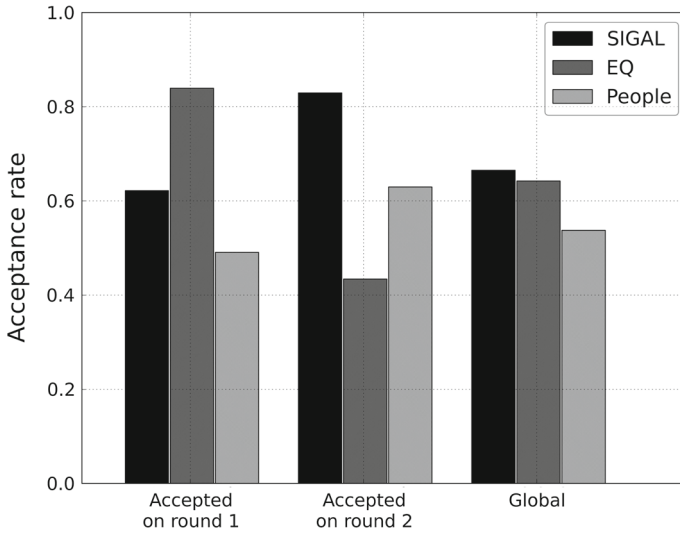


**Fig. 2** Performance comparison
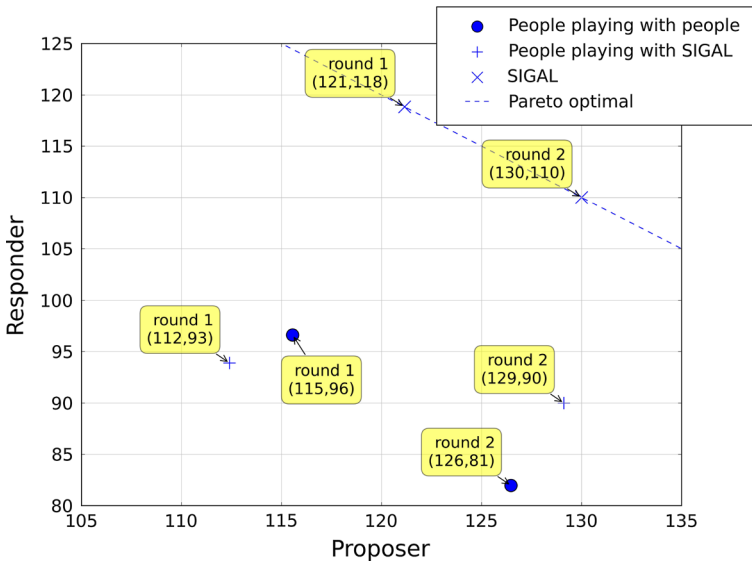
**Fig. 3** Agreement comparison



**Fig. 4** Average proposed benefit in first and second rounds

To explain the success behind SIGAL's strategy, we present a comparison of the benefit from proposals made by the SIGAL agent and people in both game rounds in Fig. 4. As shown by the Figure, both people and SIGAL made offers that were beneficial to both players in rounds 1 and 2. However, SIGAL made offers that were significantly more generous to human responders than did human proposers (118 benefit points provided by SIGAL as proposer in round 1 versus 96 points provided by human proposers; 110 benefit points provided by SIGAL as proposer in round 2 versus 81 points provided by human proposers). In fact, the proposals made by SIGAL pareto dominated the proposals made by people in both

**Table 1** Features coefficients weights

| Feature | Value |
|---|---|
| $REV_j^0$ | 0.258 |
| $REV_i^0$ | 0.035 |
| $BEN_j^n$ | 0.956 |
| $AI_i^n$ | −0.792 |
| $P.BEN_j^n$ | 0.496 |
| $P.BEN_i^n$ | 0.334 |
| Free parameter | 0.608 |

rounds 1 and 2. Thus, SIGAL learned to make offers that were better for human responders without compromising its own utility

SIGAL's strategy is highlighted by examining the weights learned for the different features of how people accept offers. As shown in Table 1, the largest weight was assigned to $BEN_j^n$, the benefit to the responder from an offer. In addition, the weight for $AI_i^n$ measuring the difference between the benefit for the proposer and responder was large and negative. This means that responders prefer proposals that provide them with large benefits, and are also competitive, in that they dislike offers that provide more to proposers than to responders. The offers made by SIGAL reflect these criteria. In particular, proposers asked more for themselves than for responders in both rounds. In contrast, SIGAL equalized the difference in benefit between proposers and responders in round 1, and decreased the difference between its own benefit and responder's benefit in round 2 as compared to human proposer.

## 8.2 Analysis: revelation of goals

We now turn to analyzing the affect of goal revelation on the behavior of SIGAL. Recall that $E_a\left(\phi_a^* = t_a\right)$ denotes the value of the game for SIGAL when deciding to reveal its goal in round 0, and behaving optimally according to its model of how people make offers. Similarly, $E_a\left(\phi_a^* = null\right)$ denotes the value of the game for SIGAL when deciding not to reveal its goal in round 0. Our model predicted no significant difference in value to SIGAL between revealing and not revealing its goal, i.e. $E_a\left(\phi_a^* = null\right) \approx E_a\left(\phi_a^* = t_a\right)$ for each type $t_a \in T_a$. Therefore we used two types of SIGAL agents, one that consistently revealed its goal at the onset of the game and one that did not reveal. In all other respects these agents followed the model described in Sect. 5. The empirical results confirmed the model's prediction, in that there was no significant difference in the performance of the two SIGAL agents for all boards and types used in the empirical study.

However, the decision of the person to reveal or not reveal its goal had a significant affect on the negotiation strategy of SIGAL. When people revealed their goals, SIGAL was significantly more competitive as compared to the case in which people did not reveal their goals. Specifically, the competitive weight $AI_i^n$ from proposals made by SIGAL in round 1 was 14 points when people revealed their goal, and −3 points when people did not reveal their goal. This means that SIGAL learned to make offers that were significantly more competitive when people revealed their goals.

## 8.3 Analysis: people's strategies

We now turn to describe several features in people's behavior demonstrating strategic behavior in the game that resembled the way SIGAL played. In the asymmetric game, people in weak-

type roles were more likely to reveal their goals than people in strong-type roles. In particular, in 65 % of games played in the asymmetric game, strong-type players engaged in "bluffing", that is, not disclosing their true goal and asking for more chips than they actually need to get to the goal in round 1. Similarly, in 56 % of games in the asymmetric condition, weak-type players engaged in bluffing. Interestingly, this trend was also apparent for the symmetric game, in which bluffing occurred in 63 % of the games played. Second, people's proposals were significantly more selfish and less generous in round 2 than in round 1. This is apparent from Fig. 2. Interestingly, people's performance was significantly higher when playing with SIGAL than when playing with other people. Specifically, people playing with SIGAL achieved an average performance of 112 points, while people playing with other people achieved an average performance of 91 points. This shows that SIGAL had a positive affect on people's play.

The last part of the analysis compares performance of subjects enlisted using Amazon Turk (MTurk) with students. There are several works dupicating lab studies using MTurk showing similar patterns of behavior [1]. In our work, we found that in general, students were better performers than MTurk workers (average score was 123 points for students and 109 points for MTurk workers). This can be explained by the fact that students were significantly more selfish as second proposers (average 45 points for students and 20 points for MTurk workers). Also, MTurk workers accepted significantly more proposals than did students in the first round (75 % average acceptance rate for MTurk compared to 59 % average acceptance rate for students). These results suggest that integrating information revelation with negotiation required more cognitive effort from people than canonical decision-making tasks.

## 8.4 Limitations

We conclude this section with describing several limitations of our approach. First, our definition of revelation games, which the SIGAL decision-making model is tailored to support, defines a single revelation phase followed by two rounds of take-it-or-leave-it offers. This model fails to describe more involved negotiations which may include additional rounds or multiple instances of revelation. Indeed, some agents may wish to reveal their goal only after not succeeding to reach agreement in several rounds. Allowing for more involved revelation protocols raises significant computational challenges which remain outside the scope of this paper. Second, the decision-making model used by SIGAL in this paper assumes that all people use tit-for-tat reasoning when predicting their performance, which may not hold in practice. Indeed, or results showed considerable variance in people's behavior for different countries. We hypothesize that learning separate decision-making models for each conuntry may additionally improve the prediction power of SIGAL and consequently, its performance.

## 9 Conclusion and future work

This paper presented an agent-design for interacting with people in "revelation games", in which participants are given the choice to truthfully reveal private information prior to negotiation. The decision-making model used by the agent reasoned about the social factors that affect people's decisions whether to reveal their goals, as well as the effects of people's revelation decisions on their negotiation behavior. The parameters of the model were estimated from data consisting of people's interaction with other people. In empirical investigations, the agent was able to outperform people playing other people as well as agents playing equilibrium strategies and was able to reach agreement significantly more often than did people.

We are currently extending this work in two directions. First, we are considering more elaborate settings in which players are able to control the extent to which they reveal their goals. Second, we are using this work as the basis for a more broad argumentation in which agents integrate explanations and justifications within their negotiation process. Lastly, information revelation strategies can also be adopted by the negotiation community at large, for example in the International Automated Negotiation Agent Competition (ANAC) [3].

## 10 Appendix

Before proving Proposition 1, we will postulate the following lemma:

**Lemma 1** *Suppose that player $j$ uses response strategy $a_j^2(\omega, t_j)$ (Eq. 2) to accept proposals $\omega$ at round 2. Then player $i$ is always indifferent between revealing and not revealing its type, regardless of the revelation decision of player $j$.*

*Proof* We will show that the the revelation decisions of player $i$ cannot affect its outcome in the game. The strategy $a_j^2(\omega, t_j)$ of player $j$ to accept the proposal by $i$ at round 2 does not depend on $i$'s type. Hence the revelation decision of $i$ in round 0 does not affect its benefit from proposals in round 2. Suppose that player $i$ revealed its type. If it can get a more beneficial proposal from player $j$ by not revealing its type, then it will reject $j$'s offer and make the more beneficial proposal in round 2. A similar argument can be used to show that player $i$ can get its more preferred outcome when it does not reveal its type. Hence the revelation decision of $i$ does not affect its outcome.

We provide a proof to Proposition 1 by showing that the strategies specified in Sect. 4.1 correspond to a perfect Bayesian equilibrium given the sufficient conditions.

*Proof* We denote the possible types for $j$ as $t_{j_1}$ and $t_{j_2}$ and the possible types for $i$ as $t_{i_1}$ and $t_{i_2}$. We prove that when one type of player $j$ is preferable to the other type (Eq. 6), there always exists a revelation equilibrium in which players reveal their goals for both types.[10] In round 2, player $i$ makes the proposal $\omega_{i,j}^{2*}(t_j)$ using Eq. 4. This proposal gives the maximum benefit for player $i$ among the proposals that are accepted by player $j$. Any proposal with higher benefit for $i$ will be rejected by player $j$.[11] Player $j$ also cannot benefit from deviation as rejecting the proposal incurs zero benefit to both players.

We proceed to round 1. We first specify why the belief update rules are part of a perfect Bayesian equilibrium. Suppose $j$ is of type $t_{j_2}$ and that this type is preferable for $j$ (without loss of generality). If $j$ revealed its type, then $i$ uses Bayes rule to assign probability 1 to type $t_{j_2}$, as specified in equilibrium. If $j$ does not reveal its type, then the resulting information set is off the equilibrium path. In this case $i$ makes an arbitrary belief update to assign probability 0 to type $t_{j_2}$. Using a similar argument, we can show that the belief update rules are part of a perfect Bayesian equilibrium when type $t_{j_1}$ is preferable to $j$. Suppose that player $i$ is of

---

[10] The existence of a preferable type for a player is a reasonable assumption in negotiation. Only if it is the case that both players prefer the proposal that is associated with the type of the other, there is no preferable type.

[11] In the case where there are no proposals that are accepted by player $j$, the benefit for player $i$ is zero.

type $t_{i_1}$. If $i$ reveals its type, then $j$ uses Bayes rule to assign probability 1 to $t_{i_2}$, as specified in equilibrium. If $i$ does not reveal its type, then the resulting information set is off the equilibrium path. In this case $j$ makes an arbitrary belief update to assign probability 1 to $t_{i_1}$. Using a similar argument, we can show that the belief update rules are part of a perfect Bayesian equilibrium when $i$ is of type $t_{i_2}$.

Next, we specify why players' actions are in equilibrium in round 1. Player $j$ makes the proposal $w_{j,i}^{1*}(t_i)$ using Eq. 5. Any proposal with a higher benefit to $j$ will necessarily be rejected by player $i$. As a result, player $j$'s proposal maximizes its benefit given $i$'s response strategy. Player $i$ also cannot benefit from deviation because player $j$'s proposal gives it the same benefit as its counter-proposal $\omega_{i,j}^{2*}(t_j)$.

Finally, we consider round 0, in which both players $i$ and $j$ reveal their types for both of their types. A deviation for player $j$ of type $t_{j_2}$ in round 0 means that $j$ chooses not to reveal its type. According to the belief-update rules, player $i$ will update its beliefs to assign probability 1 to type $t_{j_1}$. As a result, it will make a proposal in round 2 to $j$ that is less preferable to the one it would receive if it revealed its type. Thus, deviation from its revelation decision is not strictly beneficial for $j$ when it is of type $t_{j_2}$. In a similar fashion, we can show that deviation from a revelation decision is not beneficial for $j$ when it is of type $t_{j_1}$. Thus the revelation strategies of round 0 are in Equilibrium for $j$.

Due to Lemma 1, a decision by player $i$ to reveal its type in round 0 for each of its types is in equilibrium. Thus there always exists a revelation equilibrium in which players reveal their goals for both types.

We provide a proof to Proposition 1 by showing that the strategies specified in Sect. 4.2 specify a perfect Bayesian Equilibrium.

*Proof* We use backward induction. In Round 2 there are two cases. On the equilibrium path. In this case the players did not update their beliefs according to the equilibrium specification. Player $i$ makes the proposal $\omega_{i,j}^{2*}(h^1, t_i)$ using Eq. 10. This proposal gives the maximum expected benefit for player $i$, so it cannot on expectation benefit from deviation. Player $j$ also cannot benefit from deviation as rejecting the proposal incurs zero benefit to both players.

Lets examine the cases where the information set is off the equilibrium path. The first case is related to round 0: if $i$ revealed its type, this will not affect the proposal at round 2 (Proposition 2). If $j$ revealed its type, then $i$ knows its true type, and the strategies are in equilibrium as in the revelation equilibrium. The other case is related to round 1: In this case, as defined in the Equilibrium, the players do not change their beliefs, and the strategies in round 2 are as in the equilibrium. We now proceed to Round 1. we show neither player has incentive to deviate from its strategy. Suppose player $j$ of type $t_j$ wishes to deviate and propose a different proposal. If the proposal is worse off for player $i$, given its type $t_i$, then player $i$ will reject the offer and player $j$ will not benefit. If player $i$ accepts the proposal (according to Eq. 12), this means that it is greater or equal in benefit to $i$ than the equilibrium proposal in round 2. By the second condition (Eq. 14), this means that the proposal will be at most equal to the benefit that player $j$ will receive from the equilibrium proposal. So player $j$ will not benefit from making this proposal. For player $i$, it cannot benefit from rejecting $j$'s proposal. According to Eq. 12, player $i$ accepts player $j$'s proposal only if it cannot get a higher benefit from its own counter-proposal.

Finally, Lets examine if the player can benefit from deviation in Round 0. Suppose that player $j$ of type $t_{j_1}$ deviates and reveals its type. By definition, then player $i$ will update its beliefs to its true type. In this case, player $i$ will make the proposal $\omega_{i,j}^{2*}(t_{j_1})$. However, by the first condition (Eq. 13), this proposal is at most equal to the proposal that player $j$

will get from not revealing its type. Therefore it is not beneficial for player $j$ to deviate. By Proposition 1, player $i$ is indifferent between its two types. Thus the strategy for both players satisfies a perfect Bayesian equilibrium.

We conclude by showing that the sufficient conditions of proposal 2 hold in the symmetric (Fig. 1a) and the asymmetric (Fig. 1b) boards described in the paper:

Symmetric board: For the first condition (13), $\pi_j(\omega_{i,jl}^{2*}(t_j) \mid t_l) = 20$ for $l \in 1, 2$, is smaller than $\pi_j(\omega_{i,j}^{2*}(h^1, t_i) = 25$, thus the condition holds. For the second condition (14), $\omega_{i,j}^{2*}(h^1, t_i)$ let player $i$ reach its goal, and so for both types of player $j$. In such case, the only way for a player to get more points is by asking for more chips, which means that the game turns to be a zero-sum game. In such case, there is no proposal that can be more beneficial for both players, thus the condition holds.

Asymmetric board: For the first condition (13), $\pi_j(\omega_{i,jl}^{2*}(t_j) \mid t_1) = 30$ for $t_1$ as the weak type, and $\pi_j(\omega_{i,jl}^{2*}(t_j) \mid t_1) = 10$ for $t_2$ as the strong type. $\pi_j(\omega_{i,j}^{2*}(h^1, t_i)) = 35$ for $t_1$ and 15 for $t_2$, thus the condition holds. For the second condition (14), like in the symmetric board $\omega_{i,j}^{2*}(h^1, t_i)$ let both types of player $j$ and player $i$ to reach their goal, so the game turns to be a zero-sum game, and the condition holds.

## References

1. Amir, O., Rand, D. G., et al. (2012). Economic games on the internet: The effect of $1 stakes. *PloS ONE*, *7*(2), e31461.
2. Azaria, A., Rabinovich, Z., Kraus, S., & Goldman, C. V. (2011). Strategic information disclosure to people with multiple alternatives. In *Proceedings of AAAI*.
3. Baarslag, T., Fujita, K., Gerding, E. H., Hindriks, K., Ito, T., Jennings, N. R., et al. (2013). Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. *Artificial Intelligence*, *198*, 73–103.
4. Banks, J., Camerer, C. F., & Porter, D. (1994). Experimental tests of nash refinements in signaling games. *Games and Economic Behavior*, *6*, 1–31.
5. Bench-Capon, T., Atkinson, K., & McBurney, P. (2009). Altruism and agents: an argumentation based approach to designing agent decision mechanisms. In *Proceedings of AAMAS*.
6. Coehoorn, R. M., & Jennings, N. R. (2004). Learning on opponent's preferences to make effective multi-issue negotiation trade-offs. In *Proceedings of EC*.
7. de Melo, C., Carnevale, P., Read, S., Antos, D., & Gratch, J. (2012). Bayesian model of the social effects of emotion in decision-making in multiagent systems. In *Proceedings of AAMAS*.
8. Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, *88*(4), 848–881.
9. Gal, Y., Grosz, B., Kraus, S., Pfeffer, A., & Shieber, S. (2010). Agent decision-making in open-mixed networks. *Artificial Intelligence*, *174*(18), 1460–1480.
10. Gal, Y., Kraus, S., Gelfand, M., Khashan, H., & Salmon, E. (2012). Negotiating with people across cultures using an adaptive agent. *ACM Transactions on Intelligent Systems and Technology*, *3*(1).
11. Gal, Y., & Pfeffer, A. (2007). Modeling reciprocity in human bilateral negotiation. In *Proceedings of AAAI*.
12. Grosz, B., Kraus, S., Talman, S., & Stossel, B. (2004). The influence of social dependencies on Decision-Making. In *Proceedings of AAMAS*.
13. Haim, G., Gal, Y., Kraus, S., & Gelfand, M. J. (2012). A cultural sensitive agent for Human–Computer negotiation. In *Proceedings of AAMAS*.
14. Hajaj, C., Hazon, N., Sarne, D., & Elmalech, A. (2013). Search more, disclose less. In *Proceedings of AAAI*.
15. Hindriks, K., & Tykhonov, D. (2008). Opponent modelling in automated multi-issue negotiation using bayesian learning. In *Proceedings of AAMAS*, (pp. 331–338).

16. Jonker, C. M., Robu, V., & Treur, J. (2007). An agent architecture for multi-attribute negotiation using incomplete preference information. *Autonomous Agents and Multi-Agent Systems*, *15*(2), 221–252.
17. Lin, R., & Kraus, S. (2010). Can automated agents proficiently negotiate with humans? *Communications of the ACM*, *53*(1), 78–88.
18. McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society*, *60*(4), 803–836.
19. Osborne, M. J., & Rubinstein, A. (1999). *A course in game theory*. Cambridge, MA: MIT Press.
20. Oshrat, Y., Lin, R., & Kraus, S. (2009). Facing the challenge of human-agent negotiations via effective general opponent modeling. In *Proceedings of AAMAS*.
21. Pasquier, P., Hollands, R., Dignum, F., Rahwan, I., & Sonenberg, L. (2007). An empirical study of interest-based negotiation. In *Proceedings of ICEC*.
22. Peled, N., Gal, Y., & Kraus, S. (2011). A study of computational and human strategies in revelation games. In *Proceedings of AAMAS*.
23. Rosenfeld, A., & Kraus, S. (2012). Modeling agents based on aspiration adaptation theory. *Autonomous Agents and Multi-Agent Systems*, *24*(2), 221–254.
24. Sarne, D., Elmalech, A., Grosz, B. J., & Geva, M. (2011). Less is more: Restructuring decisions to improve agent search. In *Proceedings of AAMAS*.
25. Spence, A. M. (1974). *Market signaling*. Cambridge, MA: Harvard University Press.
26. Sycara, K. P. (1990). Persuasive argumentation in negotiation. *Theory and decision*, *28*(3), 203–242.
27. Wedekind, C., & Milinski, M. (1996). Human cooperation in the simultaneous and the alternating prisoner's dilemma. *PNAS*, *93*(7), 2686.