A Study of Computational and Human Strategies in Revelation Games

Noam Peled
Gonda Brain Research Center
Bar Ilan University, Israel
noam.peled@live.biu.ac.il

Ya'akov (Kobi) Gal
Dept. of Information Systems Engineering
Ben-Gurion University, Israel
kobig@bgu.ac.il

Sarit Kraus *
Dept. of Computer Science
Bar Ilan University, Israel
sarit@cs.biu.ac.il

ABSTRACT

Revelation games are bilateral bargaining games in which agents may choose to truthfully reveal their private information before engaging in multiple rounds of negotiation. They are analogous to real-world situations in which people need to decide whether to disclose information such as medical records or university transcripts when negotiating over health plans and business transactions. This paper presents an agent-design that is able to negotiate proficiently with people in a revelation game with different dependencies that hold between players. The agent modeled the social factors that affect the players' revelation decisions on people's negotiation behavior. It was empirically shown to outperform people in empirical evaluations as well as agents playing equilibrium strategies. It was also more likely to reach agreement than people or equilibrium agents.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]

General Terms

Experimentation

Keywords

Human-robot/agent interaction, Negotiation

1. INTRODUCTION

In many negotiation settings, participants lack information about each other's preferences, often hindering their ability to reach beneficial agreements. This paper presents a study of a particular class of such settings we call "revelation games". In these settings, 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: Signaling games [13], in which players choose whether to convey private information to each other, and bargaining [10], in which players engage in multiple negotiation rounds. Revelation games are analogous to real-world scenarios in which parties may choose to truthfully reveal information before negotiation ensues. For example, consider a scenario in which company employees negotiate over the conditions of their employer-sponsored health insurance policy. The employees can wave the right to keep their medical records private. The disclosure of this information to the employer is necessarily truthful and is not associated with a cost to the employees. It may provide employees with favorable conditions when negotiating over future health policies. However, many people choose not to disclose medical records to their employees, fearing they may be compromised by this information.

This paper describes a new agent design that uses a decision-theoretic approach to negotiate proficiently with people in revelation games. The agent explicitly reasons about the social factors that affect people's decisions whether to reveal private information, as well as the effects of people's revelation decisions on their negotiation behavior. It combines a prediction model of people's behavior in the game with a decision-theoretic approach to make optimal decisions. The parameters of this model were estimated from data consisting of human play. The agent was evaluated playing new people and an agent playing equilibrium strategies in a revelation game that varied the dependency relationships between players. The results showed that the agent was able to outperform human players 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.

The contributions of this paper are fourfold. First, it formally presents revelation games as a new type of interaction which supports controlled revelation of private information. Second, it presents a model of human behavior that explicitly reasons about the social factors that affect people's negotiation behavior as well as the effects of players' revelation decisions on people's negotiation behavior. Third, it incorporates this model into a decision-making paradigm for an agent that uses the model to make optimal decisions in reve-

*Sarit Kraus is also affiliated with the University of Maryland Institute for Advanced Computer Studies.

Cite as: A Study of Computational and Human Strategies in Revelation Games, Peled N, Gal Y and Kraus S, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. XXX-XXX.

lation games. Lastly, it provides an empirical analysis of this agent, showing that the agent is able to outperform people as well as more likely to reach agreement than people.

2. RELATED WORK

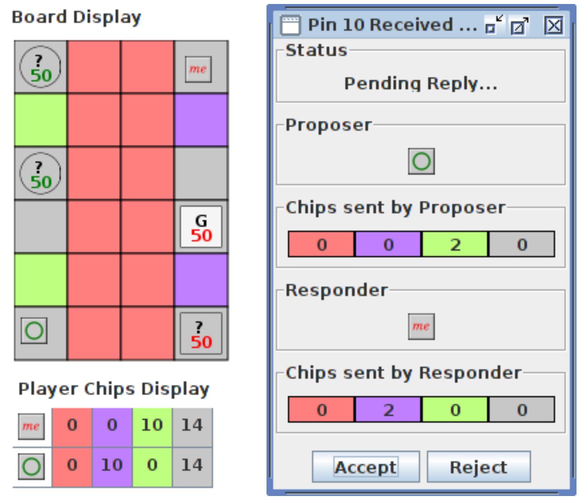
Our work is related to studies in AI that use opponent modeling to build agents for repeated negotiation in heterogeneous human-computer settings. These include the KBAgent that made offers with multiple attributes in settings which supported opting out options, and partial agreements [11]. This agent 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 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 [6] or approximation heuristics [7] to estimate people's preferences in negotiation and integrated this model with a predefined concession strategy to make offers. Bench-Capon [2] provide an argumentation based mechanism for explaining human behavior in the ultimatum game. We extend these works in two ways, first in developing a partially strategic model of people's negotiation behavior and second in formalizing an optimal decision-making paradigm for agents using this model. Gal and Pfeffer [4] 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. 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. Most of these works assume that agents follow pre-defined strategies for revealing information [12, 14] 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 [1]. On the other hand, people's bargaining behavior does not adhere to equilibrium [3, 9], and computers cannot use such strategies to negotiate well with people [8]. 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 [5], 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 goal 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



(a) (b)

Figure 1: (a) A Colored Trails revelation game shown from Bob's point of view. (b) Bob's offer

each player could only see its own goal. An example of a CT revelation game is shown in Figure 1. Here, the "me" and "O" icons represent two players, Bob and Alice, respectively. Each player has two possible goals. Bob's true goal is located three steps below the "me" icon (appearing as a white G square). Bob's other goal is located two steps below his true goal (appearing as a grey "?" square). Alice's possible goals are presented as two grey "?" circles, located three and five steps above Alice's "O" icon. The board is presented from Bob's point of view. Bob can see its true goal location but Alice does not observe it. 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.

Our CT game progresses in three phases with associated time limits. In the revelation phase, both players can choose to truthfully reveal their goal to the other player.[1] In the proposal phase, 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. 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] Suppose for example that Alice's true goal is five steps above the position of her icon (Bob does not see this goal if Alice does not reveal it). Bob is missing one chip to get to the goal while Alice is missing two chips; the score for Alice is 70 points and for

[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.

Bob is 90 points.

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, if there is a second round, the proposer in this round has an advantage, 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. THE SIGAL AGENT

The Sigmoid Acceptance Learning Agent (SIGAL) developed for this study 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. Each player has a type t_i that represents the true position of its goal on the board.[3] Let ω^n represent an offer ω made by a proposer player at round $n \in \{1, 2\}$ in a game. Let $r^n \in \{accept, reject\}$ represent the response to ω^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 $s_i(\emptyset)$.[4] We denote the benefit to player i from ω^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 ω^n and the NNA score:

$$\pi_i(\omega^n \mid t_i) = s_i(\omega^n \mid t_i) - s_i(\emptyset) \quad (1)$$

Let T_i denote a set of types for player i. Let ϕ_i denote player i's decision whether to reveal its type at the onset of the game, which we will refer to as round 0. Let $\Phi_i = t_i^k$ denote the event in which i reveals its type $t_i^k \in T_i$, and let $\Phi_i = \mathsf{null}$ 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. We define h^0 and h^1 as follows:

$$h^0 = \{\phi_i, \phi_j\}; h^1 = \{h^0, \omega^1, r^1\} \quad (2)$$

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

We now specify the strategy of SIGAL for the revelation game defined in Section 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 Section 5. We begin by describing the negotiation strategies of SIGAL for rounds 2 and 1.

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

[4]Note that 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. 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.

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

$$\begin{aligned} &E_a\left(\omega^2_{a,p} \mid h^1, t_a\right) = \\ &\quad \pi_a(\omega^2_{a,p} \mid t_a) \cdot p(r^2_p = accept \mid \omega^2_{a,p}, h^1) + \\ &\quad \pi_a(\emptyset \mid t_a) \cdot p(r^2_p = reject \mid \omega^2_{a,p}, h^1) \end{aligned} \quad (3)$$

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.

$$\omega^{2*}_{a,p} = argmax_{\omega^2_{a,p}} E_a\left(\omega^2_{a,p} \mid h^1, t_a\right) \quad (4)$$

If SIGAL is the second responder, its optimal action is to accept any proposal from the person that gives it positive benefit. Let $r^{2*}_a(\omega^2_{p,a} \mid h^1)$ denote the response of SIGAL to offer $\omega^2_{p,a}$, defined as

$$r^{2*}_a(\omega^2_{p,a} \mid h^1) = \begin{cases} accept & \pi_a(\omega^2_{p,a} \mid t_a) > 0 \\ reject & otherwise \end{cases} \quad (5)$$

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

$$\begin{aligned} &\pi_a\left(r^{2*}_a \mid \omega^2_{p,a}, h^1, t_a\right) = \\ &\quad \begin{cases} \pi_a(\omega^2_{p,a} \mid t_a) & r^{2*}_a(\omega^2_{p,a} \mid h^1) = accept \\ \pi_a(\emptyset \mid t_a) & otherwise \end{cases} \end{aligned} \quad (6)$$

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

$$\begin{aligned} &E_a(resp^2 \mid h^1, t_a) = \\ &\quad \sum_{\omega^2_{p,a}} p(\omega^2_{p,a} \mid h^1) \cdot \pi_a(r^{2*}_a \mid \omega^2_{p,a}, h^1, t_a) \end{aligned} \quad (7)$$

where $\pi_a(r^{2*}_a \mid \omega^2_{p,a}, h^1, t_a)$ is defined in Equation 6.

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

$$\begin{aligned} &E_a\left(\omega^1_{a,p} \mid h^0, t_a\right) = \\ &\quad \pi_a(\omega^1_{a,p} \mid t_a) \cdot p(r^1_p = accept \mid \omega^1_{a,p}, h^0) + \\ &\quad E_a(resp^2 \mid h^1, t_a) \cdot p(r^1_p = reject \mid \omega^1_{a,p}, h^0) \end{aligned} \quad (8)$$

Where $h^1 = \{h^0, \omega^1_{a,p}, r^1_p = reject\}$. SIGAL will propose an offer in round 1 that maximizes its expected benefit in this round:

$$\omega^{1*}_{a,p} = argmax_{\omega^1_{a,p}} E_a\left(\omega^1_{a,p} \mid h^0, t_a\right) \quad (9)$$

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 $\omega^{2*}_{a,p}$ in round 2, given its model of how people respond to offers in round 2:

$$r_a^{1*}(\omega_{p,a}^1 \mid h^0) = \begin{cases} accept & \pi_a(\omega_{p,a}^1 \mid t_a) > \\ & E_a\left(\omega_{a,p}^{2*} \mid h^1, t_a\right) \\ reject & otherwise \end{cases} \quad (10)$$

Here, $h^1 = \{h^0, \omega_{p,a}^1, r_a^1 = reject\}$, $\pi_a(\omega_{p,a}^1 \mid t_a)$ is defined in Equation 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 Equation 3.

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:

$$\pi_a\left(r_a^{1*} \mid \omega_{p,a}^1, h^0, t_a\right) = \begin{cases} \pi_a(\omega_{p,a}^1 \mid t_a) & r_a^{1*}(\omega_{p,a}^1 \mid h^0) = accept \\ E_a\left(\omega_{a,p}^{2*} \mid h^1, t_a\right) & otherwise \end{cases} \quad (11)$$

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.

$$E_a\left(resp^1 \mid h^0, t_a\right) = \sum_{t_p \in T_p} p(t_p \mid h^0) \cdot \left(\sum_{\omega_{p,a}^1} p(\omega_{p,a}^1 \mid t_p, h^0) \cdot \pi_a\left(r_a^{1*} \mid \omega_{p,a}^1, h^0, t_a\right)\right) \quad (12)$$

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:

$$E_a(h^0, t_a) = p(resp) \cdot E_a\left(resp^1 \mid h^0, t_a\right) + p(prop) \cdot E_a\left(\omega_{a,p}^{1*} \mid h^0, t_a\right) \quad (13)$$

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 Equation 8.

Because players do not observe each other's revelation decisions, the expected benefit for a revelation decision ϕ_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 $p(\phi_p = t_p)$ 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.

$$E_a\left(\phi_a\right) = \sum_{t_p \in T_p} [p(\phi_p = t_p) \cdot E_a\left(h^0 = \{\phi_a, \phi_p = t_p\}, t_a\right) + p(\phi_p = null) \cdot E_a\left(h^0 = \{\phi_a, \phi_p = null\}, t_a\right)] \quad (14)$$

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

$$\phi_a^* = \begin{cases} t_a & E_a\left(\phi_a = t_a\right) \geq \\ & E_a\left(\phi_a = null\right) \\ null & otherwise \end{cases} \quad (15)$$

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 ω^n, which modified Equation 1.

$$\pi'_a(\omega^n \mid t_a) = \frac{\pi_a(\omega^n \mid t_a)^{(1-\rho)}}{1-\rho} \quad (16)$$

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

5. 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.

5.1 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 Equation 1. Otherwise, the value of this feature is the expected

[5] Both of these patterns were confirmed empirically, as shown in the Results section.

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 $\pi_i(\omega_{i,j}^n, accept \mid t_i)$ 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 $\pi_i(\omega_{i,j}^1, accept \mid t_i)$ if $n = 2$ and 0 otherwise.

To illustrate, consider the CT board game shown in Figure 1. 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 $P.BEN_i^2 = 115$.

5.1.1 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 ω^n and history h^{n-1} to an (ordered) set of feature values x^n as follows.[6]

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

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 [4]. We define the probability of acceptance for a particular features values x^n by a responder to be

$$p(r_i^n = accept \mid \omega^n, h^{n-1}) = \frac{1}{1 + e^{-u(x^n)}} \qquad (17)$$

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.

5.1.2 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 $p(r_p^n = accept \mid \omega_{a,p}^n, h^{n-1})$, as follows:

$$\sum_{\omega_{i,p}^n \in \Omega_{i,p}} \left(p(r_j^n = accept \mid \omega_{i,p}^n, h^{n-1}) - y(r_j^n \mid \omega_{i,p}^n)\right)^2 \qquad (18)$$

where $p(r_j^n = accept \mid \omega_{i,j}^n, h^{n-1})$ is defined in Equation 17.

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 overfitting 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

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

[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 Empirical Methodology Section.

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 ϕ_a^* is defined in Equation 15. 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.

5.2 Proposing and Revealing

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

5.2.1 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 $p(\omega_{p,j}^1 \mid t_p, \phi_i)$ denote the probability that a human proposer of type t_p offers $\omega_{p,j}^1$ in round 1. Let $N_{t_p,\phi_p}(\pi_p(\omega_{p,j}^1 \mid t_p))$ denote the number of first proposals which gives the human a benefit of $\pi_p(\omega_{p,j}^1 \mid t_p)$, given the human is of type t_p and its revelation decision was ϕ_p. Let $N_{t_p,\phi_p}(\Omega_{p,j}^1)$ denote the number of the first proposal in this subset. $p(\omega_{p,j}^1 \mid t_p, \phi_p)$ is defined as:

$$p(\omega_{p,j}^1 \mid t_p, \phi_p) = \frac{N_{t_p,\phi_p}(\pi_p(\omega_{p,j}^1 \mid t_p))}{N_{t_p,\phi_p}(\Omega_{p,j}^1)} \tag{19}$$

5.2.2 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 much sparse to calculate it directly as described in Subsection 5.2.1. Inspired by studies showing that people engage in tit-for-tat reasoning [15] we used this principal to model the counter-proposals made by people. We assumed that a responder player i will be proposed offer $\omega_{p,i}^2$ by a human player in the second round with benefit $\pi_i(\omega_{p,i}^2 \mid t_i)$ that is equal to the benefit $\pi_p(\omega_{i,p}^1 \mid t_p)$ from offer $\omega_{i,p}^1$ made to the people 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 $N_{\Omega_{p,i}^2}(\pi_p(\omega_{i,p}^1 \mid t_p))$ denote the number of counter-proposals $\omega_{p,i}^2$ which give benefit $\pi_p(\omega_{i,p}^1 \mid t_p)$. We assume that there always exists at least one proposal that meets this criterion, i.e., $N_{\Omega_{p,i}^2}(\pi_p(\omega_{i,p}^1 \mid t_p)) \neq 0$. The "tit for tat" heuristic is as follows:

$$p(\omega_{p,i}^2 \mid h^1) = \begin{cases} 0 & \pi_i(\omega_{p,i}^2) \neq \pi_p(\omega_{i,p}^1) \\ 1/N_{\Omega_{p,i}^2}(\pi_p(\omega_{i,p}^1 \mid t_P)) & otherwise \end{cases} \tag{20}$$

This heuristic is used in Equation 7 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:

$$p(\phi_p \mid t_p) = \frac{N_{t_p}(\phi_p)}{N_{t_p}} \tag{21}$$

6. 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 228 students enrolled in a computer science or software engineering program at several universities and colleges. 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 two boards in the study. In the "asymmetric board", one of the players needed a single chip of a particular color to reach its goal, while the other player needed two chips of a particular color to reach its respective goal. This is the board that is shown in Figure 1. We also used a "symmetric board" in which both players needed a single chip of one of two possible colors to get to their goal.

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 the board layout, and 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 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

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

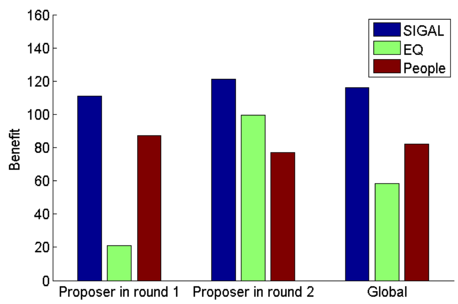


Figure 2: Performance comparison

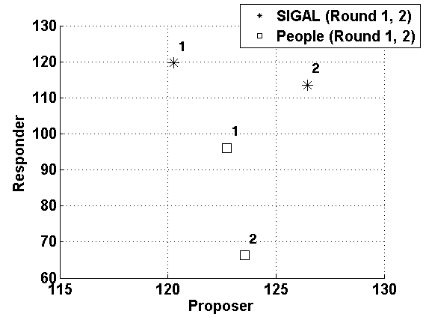


Figure 3: Average Proposed Benefit in First and Second rounds

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 (110 games). The fourth pool (118 games) consisted of people playing an agent using an equilibrium strategy to play revelation games.

7. RESULTS AND DISCUSSION

The performance of SIGAL was measured by comparing its performance against people in the third pool with people's play in the remaining third of the first pool.[10] 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.

7.1 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-negation alternative score) for different roles (proposers and responder). As shown by the figure, the SIGAL agent outperformed people in all roles (111 points as proposer in round 1 versus 87 points for human proposers in round 1; 121 points as proposer in round 2 versus 77 points for human proposers in round 2).

The SIGAL agent was also more successful at reaching agreements than were people. Only 2% of games in which SIGAL played people did not reach agreement (in first or second round), while 27% of games in which people played other people did not reach agreement. In particular, offers made by SIGAL in round 2 were accepted 87% of the time, while offers made by people in round 2 were only accepted 14% 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 points to people and 215 to itself. Most of these offers were not accepted. 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 were accepted, but this did not aid its performance.

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 Figure 3. 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 (120 points benefit provided by SIGAL as proposer in round 1 versus 96 points provided by human proposers; 114 points benefit provided by SIGAL as proposer in round 2 versus 66 points provided by human proposers). As shown by the figure, there was no significant differences between the benefit to SIGAL from offers made by SIGAL itself and people (121 points to SIGAL versus 123 points to people for round 1 and 126 points versus 124 points in round 2). In particular, all of SIGAL's proposals enabled the responder to reach its goal. Thus, SIGAL was able to learn 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.

7.2 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

[10] Although this portion corresponds to only 22 games played by people, it was sufficient to achieve statistical significance.

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

Table 1: Features coefficients weights

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. $E_a(\phi_a^* = null) \approx E_a(\phi_a^* = t_a)$ 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 Section 4. 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. The results described in this section average over the revealing and non-revealing types of SIGAL agents.

This was confirmed by the empirical results, in which the average performance of the SIGAL agent when revealing its goal was 114 points ($n = 52$), while the average performance of SIGAL when not revealing its goal was 118 points ($n = 58$). This difference was not significantly significant in the $p < 0.05$ range.

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 learned to ask for more benefit for itself as compared to the case in which people did not reveal their goals. For example, when playing the asymmetric board, the non-revealing SIGAL agents learns to ask 125 points for itself if the person reveals its goal, and only 115 points for itself if the person did not reveal. In this case SIGAL took advantage of the fact that the type of the human responder is known, but its own type is not known.

Lastly, the probabilities that people revealed their goals, as learned from the training set, were as follows: 37.14% and 46.27% in the asymmetric board were missing one, and two chips to get to the goal, respectively, and 41.13% for the symmetric board, in which both players were only missing one chip. Interestingly, people missing two chips to get to the goal were most likely to reveal their type. We hypothesize this was to justify their request for their missing chips from the other player.

8. 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.

9. ACKNOWLEDGMENTS

This research was partially funded by European Union Seventh Framework Programme agreement no. PIRG07-GA-2010-268362, by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant number W911NF-08-1-0144 and by NSF grant 0705587. Thanks to Yael Blumberg and Ofra Amir for programming efforts and useful comments.

10. REFERENCES

[1] J. Banks, C. F. Camerer, and D. Porter. Experimental tests of nash refinements in signaling games. *Games and Economic Behavior*, 6:1–31, 1994.

[2] T. Bench-Capon, K. Atkinson, and P. McBurney. Altruism and agents: an argumentation based approach to designing agent decision mechanisms. In *Proc. of AAMAS*, 2009.

[3] I. Erev and A. E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4):848–881, 1998.

[4] Y. Gal and A. Pfeffer. Modeling reciprocity in human bilateral negotiation. In *Proc. of AAAI*, 2007.

[5] B. Grosz, S. Kraus, S. Talman, and B. Stossel. The influence of social dependencies on Decision-Making. In *Proc. of AAMAS*, 2004.

[6] K. Hindriks and D. Tykhonov. Opponent modelling in automated multi-issue negotiation using bayesian learning. In *Proc. of AAMAS*, pages 331–338, 2008.

[7] C. M. Jonker, V. Robu, and J. Treur. An agent architecture for multi-attribute negotiation using incomplete preference information. *Autonomous Agents and Multi-Agent Systems*, 15(2):221–252, 2007.

[8] R. Lin and S. Kraus. Can automated agents proficiently negotiate with humans? *Communications of the ACM*, 53(1):78–88, 2010.

[9] R. D. McKelvey and T. R. Palfrey. An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society*, 60(4):803–836, 1992.

[10] M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT press, 1999.

[11] Y. Oshrat, R. Lin, and S. Kraus. Facing the challenge of human-agent negotiations via effective general opponent modeling. In *Proc. of AAMAS*, 2009.

[12] P. Pasquier, R. Hollands, F. Dignum, I. Rahwan, and L. Sonenberg. An empirical study of interest-based negotiation. In *Proc. of EC*, 2007.

[13] A. M. Spence. *Market signaling*. Harvard Univ. Press, 1974.

[14] K. Sycara. Persuasive argumentation in negotiation. *Theory and decision*, 28(3):203–242, 1990.

[15] C. Wedekind and M. Milinski. Human cooperation in the simultaneous and the alternating prisoner's dilemma. *PNAS*, 93(7):2686, 1996.