Artificial Intelligence 246 (2017) 34–52

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Human–computer negotiation in a three player market setting

Galit Haim [a], Ya'akov (Kobi) Gal [b,*], Bo An [c], Sarit Kraus [d,e]

[a] *The College of Management Academic Studies, Israel*
[b] *Dept. of Information Systems Engineering, Ben-Gurion University, Israel*
[c] *School of Computer Science and Engineering, Nanyang Technological University, Singapore*
[d] *Dept. of Computer Science, Bar-Ilan University, Israel*
[e] *University of Maryland Institute for Advanced Computer Studies, United States*

ARTICLE INFO

Article history:
Received 16 January 2016
Received in revised form 17 January 2017
Accepted 20 January 2017
Available online 3 February 2017

Keywords:
Human–computer decision-making
Multi-agent negotiation
Empirical work

ABSTRACT

This paper proposes a novel agent-design for a three-player game involving human players and computer agents. The game is analogous to settings in which participants repeatedly negotiate over contracts, such as cell-phones and credit card plans. The game comprises three players, two service providers who compete to sign contracts with a single customer player. The service providers compete to make repeated contract offers to the customer consisting of resource exchanges in the game. Customers can join and leave contracts at will. We computed sub-game perfect equilibrium strategies for all players that were based on making contracts involving commitments between the customer player and one of the service provider players. We conducted extensive empirical studies (spanning over 500 participants) comparing the performance of computer agents using different types of equilibrium strategies with that of people in three different countries, the U.S., Israel and China, that are characterized by cultural differences in how people make contracts in the game. Two human participants played a single computer agent in various role configurations in the game. For the customer role, agents using equilibrium strategies were able to obtain a higher score than people playing the same role in three countries. For the service provider role, agents using equilibrium strategies that reasoned about possibly irrational behavior were able to obtain higher scores than people (as well as agents that did not reason about irrational behavior). This work shows that for particular market settings involving competition between service providers, equilibrium strategies can be a successful design paradigm for computer agents without relying on data driven approaches.

1. Introduction

Web-based markets are becoming increasingly prevalent in our daily lives [23,5]. Examples include cell-phone and credit card plans, car insurance, bank accounts, pension plans and many more. The growth of these market is facilitated by technology, which allows service providers to tailor plans to the needs of different types of customers, while customers are able to change plans and service providers easily and quickly.

In addition to these benefits, web-based markets also bring new challenges. From the perspective of customers, the existence of a wide array of different plans and possible service providers makes it time consuming and challenging to

* Corresponding author.
E-mail address: kobig@bgu.ac.il (Y. Gal).

http://dx.doi.org/10.1016/j.artint.2017.01.003
0004-3702/

negotiate over contracts. From the perspective of suppliers, the need to attract and continuously maintain a large base of customers makes it difficult to stay profitable.

This paper studies how computer agents can balance the trade-off between profitability and competition when negotiating over contracts with people in a three-player market settings. The first contribution of this paper is the design of a "contract game" which is analogous to a market setting in which participants need to reach agreements on contracts over time. The game comprises three players, two service providers and one customer that are allocated resources at the onset of the game. The service providers compete to make contract offers to the customer that consist of resource exchanges. The objectives for all players is to maximize their profits, measured by the amount of resources they are able to collect by the end of the game and whether they can reach the goal.

The second contribution of this paper is to formally define the type of offers and contracts that arise in the contract game and to define sub-game perfect equilibrium strategies for the game. The equilibrium explicitly reasons about the probability that a customer player makes irrational moves during the game. We show that for a particular class of contract games, the contracts reached in equilibria involve a commitment between the customer and one of the service providers that prevent the customer from signing a contract with the other provider.

The third contribution of this paper is to evaluate computer agents playing the equilibria strategies in extensive empirical studies carried out in three different countries, the U.S., Israel and China. We hypothesized the following: First, that people will vary in the way they play the game, but by including commitments in their strategies the equilibrium agents will be able to play well with people. Second, that defining off-the-equilibrium path strategies is crucial to interacting with people in the game, because they may not necessarily adhere to the specified equilibrium strategies.

We ran several configurations in which two human participants played with a single agent participant in various role configurations in the game. The computer agents used several types of equilibrium strategies that varied in whether they reason about possible irrational behavior by people. We measured whether the agents can (gain more utility) in the same role (service provider or responder) when playing other people (not when playing them directly). This supports the use of the computer agents as proxies for people who negotiate with other people or computer agents in e-commerce situations. Our results showed that the computer agent using the sub-game perfect equilibrium strategies for the customer role was able to outperform people in the same role all three countries. In particular, the customer agent made significantly more commitment type proposals than people, and requested significantly more resources from service providers than did people. The service provider agent was able to gain more utility than people playing the same role by reasoning about the probability that people will make irrational moves in the game. This probability was tuned separately for each country using prior data of people's play, which enabled the service provider to adapt different negotiation behavior for each country.

Lastly, we saw significant differences in the performance of the provider agents playing Nash equilibrium strategies between the different countries. Specifically, in China, people were able to outperform the provider agent, while in Israel the performance of the service provider agent was similar to that of people.

Our results have insight for agent-designer in human–computer negotiation settings in which it is difficult to use data-driven approaches. We show that despite people's inherent tendency to deviate from rational strategies, equilibrium strategies can form the basis of computational agents for negotiating with people in different cultures, provided that two conditions are met: First, these agents directly reason about people's possible irrational behavior in the game. Second, the agents compute and are able to use off-the-equilibrium path strategies. The paper provides further evidence that although equilibrium strategies are prone to fail in game settings such as the ultimatum game they can be successful in other domains, namely market competition settings, and that the extent to which people engage in sub-optimal behavior seems to be culturally influenced. It is the first study showing the value of equilibrium strategies for negotiating with people from different cultures.

The remainder of this paper is organized as follows. The next section describes relevant work on automated negotiations that include both people and computer agents. Section 3 presents the game that we used in our study. Our three-player market game was configured using the Colored Trails (CT) game [10] which provides a realistic analog to task settings. Section 4 provides a formalization of the game and provides necessary definitions. Section 5 presents the sub-game perfect equilibrium and correctness proof. Section 6 presents an empirical study with 500 human subjects from three countries. We end this section with a discussion on the reasons for the successes of the computer agent using the sub-game perfect equilibrium strategies playing the customer role and the modified agent that plays the role of a service provider. We characterize the situations in which sub-game perfect equilibrium agents can be beneficial and discuss the benefits and limitations of our approach. Lastly, Section 7 concludes the paper and presents some ongoing and future directions on agent-design for market settings that involve people.

2. Related work

Significant work exists on agent design for negotiation between multiple solely computational players [1, inter alia]. In the seminal work on analyzing bargaining agents' subgame perfect equilibrium strategies, Rubinstein [31] studies the alternating-offers game with infinite horizon. Many works in multi-agent systems focused on designing heuristic strategies for more complex negotiation problems where it is difficult to derive equilibrium strategies [3,6,19,25,35]. Gatti et al. [13] provided an algorithm to compute agents' equilibrium strategies in bilateral bargaining with one-sided uncertain deadlines. Sandholm and Vulkan [32] computed the sequential equilibrium for a continuous time bilateral bargaining setting in which agents don't know each other's deadlines. An et al. [2] formalized how uncertainty over deadlines and reserve prices can affect equilibrium strategies in one-to-many and many-to-many negotiation scenarios in which agents follow alternating-offers bargaining protocols with a discount factor. Sandholm and Zhou studied equilibrium in negotiation in which agents could opt out of a commitment by a penalty fee [33]. Kalandrakis [16] studied bargaining behavior among three players and formalized a Markov perfect Nash equilibrium that depends on the state of the world using a dynamic game formalism. All of the works above did not consider possibly bounded rational behavior by participants.

Our work also relates to works studying negotiation and bargaining behavior between people and computer agents in AI. Ficici and Pfeffer used machine learning to model the belief hierarchies that people use when they make decisions in one-shot interaction scenarios [8,9]. Van Wissen at el. [34] studied team formation in human–computer teams in which players negotiated over contracts. None of these works considered an agent-design for repeated negotiation with people. Other work has addressed some of the computational challenges arising in repeated negotiation between people and computer agents in two-player settings [12,17,20]. Azaria et al. [4] studied negotiation over completing a set of tasks in a crowd-sourcing environment. They implemented an agent which negotiated with people from the USA and from India. Rosenfeld et al. [30] studied how to design an effective chat agent for negotiations with people. Lastly, Peled et al. [28] used equilibrium agents to play with people in a two-round negotiation setting of incomplete information. These agents were outperformed by agents using machine learning methods that predicted how people reveal their goals during negotiation. All of these works included two participants, which significantly reduces the complexity of the decision-making. Kraus developed an agent for negotiating with people in a game that consists of seven players. However, it was tailored to a specific domain of the Diplomacy game [18].

Negotiation between people with different cultures has been studied extensively in the psychological and social science literature (Gal et al. [11] inter alia). There is also an increasing literature on using computational models of human negotiation behavior that reasons about cultural differences, although none of these works has provided an equilibrium based approach towards interacting with people. Mell et al. [24] studied the impact of trust in automated negotiation agents within American and Indian cultures. Gal et al. [11] proposed an agent for negotiating with people in the U.S. and Lebanon in a two-player CT setting in which agreements were not binding. They showed that an agent that used hand-designed rules to make and respond to offers could adapt to the different reliability and generosity measures that were exhibited by people in the different cultures. Haim et al. [14] proposed an alternative agent design for this setting that used machine learning to predict how people respond to offers and fulfill their agreements in each culture. This prediction was combined with a decision-theoretic model that was used by the agents to make and respond to offers. This agent was able to outperform people in the three aforementioned countries.

Common approaches in the literature for human–computer negotiation in bilateral settings rely on machine learning [15,14] or rule-based approaches to adapt to people's behavior at each stage in the game [11,21]. Such approaches are not feasible in our setting, in which the number of states is prohibitively large, and it is not possible to collect data apriori for every conceivable occurrence in the game.

Nouri and Traum [27] showed that it was possible to predict the cultural affiliation of a player (whether American or Indian) when playing 10 rounds of the ultimatum game with a virtual human using a baseline negotiation strategy. They were able to predict people's negotiation behavior in later rounds of the game using support vector machines. In further work, they used a multi-attribute utility function to model social factors that affect people's play in ultimatum and centipede games [26]. The parameters of the model were elicited from people who filled a social value questionnaire prior to playing the game. Mascarenhas et al. [22] used a theory of mind model to build agents that interact with students in a simulated virtual world. Their setting uses characters exhibiting emotional cues and empathy and adapts its reciprocal behavior based on the cultural affiliation of its partner.

3. Implementation: colored trails

Our three-player market setting was configured using the Colored Trails (CT) game [10]. It consists of a game that interleaves negotiation to reach agreements and decisions on whether to accept or reject an agreement, to whom to propose a proposal, and the movement strategy.

3.1. The contract game

There are 3 players, one is the customer (CS) player and two players are the service providers (SP_g and SP_y) players. The CS player moves on a board of color squares $m \times n$ grid. Fig. 1 shows a snapshot of the game from the perspective of a CS player (the "me" player). The SP_g player is designated as the square icons located at the far-left corner of the first row (on the gray goal square), and the SP_y player is designated as the oval goal icon on the far-right corner of the first row (on the yellow goal square). These two squares on the board were designated as the goal squares. The board also shows the location of the CS player icon on the last line of the board in the middle column, nine steps away from each goal square.

At the beginning of the game, each player has a set of colored chips, in which the amount and the colors of the chips may differ from one player to another. The game is divided into several rounds. Each round entails a negotiation phase between the customer and the providers, and a movement phase in which the customer player is free to utilize its chips to move on the board. In the negotiation phase, the SP players or the CS can act as a "Proposer" or as a "Responder". The

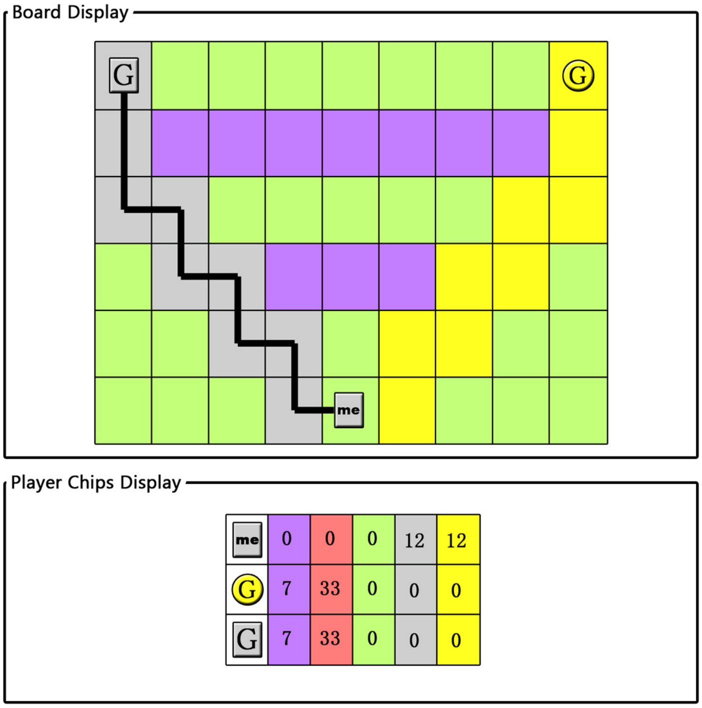


Fig. 1. Snapshot of the contract game. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

players switch their roles, such that the first proposer in the previous negotiation phase was designated as a responder in the next negotiation phase, and vice versa. When the CS is the proposer, it can send a proposal to only *one* of the providers. When the CS is the responder, each of the providers may send it a proposal. The CS player can see the proposals from both service providers while the service providers cannot see each other's proposals. Once the CS player receives a proposal, it may accept or reject the proposal, but it can accept only one such proposal in each round. Once the responder accepts a proposal, the chips are automatically exchanged between the proposer and the responder of the proposal. The movement phase is analogous to the customer player using its resources (chips) to perform individual tasks (moving on the board). The CS can choose where to move according to the chips it has, and can move any number of squares (up, right or left but not diagonally) according to the chips in its possession.

3.2. Game termination and scoring

The phases described above repeat until the game terminates, which occurs when one of the following conditions holds.

1. The CS does not move for two consecutive rounds.
2. The CS reaches one of the goal-squares belonging to one of the providers.

The players' scores are computed at an intermediate or terminal point in the game as follows:

- 150 Points to both the customer and the provider whose goal-square was reached by the customer (if any).
- 5 bonus points for any chip left in a player's possession. For example, at the beginning of the game, as shown in Fig. 1, the CS player has 24 chips and its score is 120, whereas the SPs has 40 chips each and their initial score is 200 each.

The objective of the game for the CS is to reach the goal of one of the providers, and to try to use as few chips as possible in order to end the game with a large number of chips. In this game, all players have full knowledge about the board and chips, but both providers repeatedly compete to make contracts with the customer player. The score of each player does not depend on the scores of other players.

An advantage of using CT is that it provides a realistic analog to task settings, highlighting the interaction among goals, tasks required to achieve these goals and resources needed for completing tasks. In CT, chips correspond to agent capabilities and skills required to fulfill tasks. Different squares on the board represent different types of tasks. A player's possession of a chip of a certain color corresponds to having the skill available for use at that time.

Table 1
Notations used in contract game model.

Notations	Meaning
CS	The customer player.
SP_i, SP_j	The service provider players (SP_g, SP_y).
b^*	A bonus that both the CS and the SP_i receive if the CS player reached G_i.
c	The number of points a player will get for each chip it has at the end of the game.
$u_i(s)$	The score of any player i at a non-terminal state s in the game.
(x, y)	The location of the CS player on the board.
C_{cs}, C_1, C_2	The set of chips of the customer player, SP_g and SP_y respectively.
r	The round of the game.
d	Dormant indicator.
s	The state of the game.
G_1, G_2	The goals for the customer player, SP_g and SP_y respectively.
P	A possible path from (x_1, y_1) to (x_k, y_k) on the board.
C_P	The set of *needed chips* to go through a path P.
$Res(s, P)$	The *resulting* state of CS moving according to P.
$Res(s, O)$	The resulting state from the acceptance of an offer $O = (O_{cs}, O_i)$.
O	An offer is a pair (O_{cs}, O_i) where O_{cs} and O_i is the set of chips sent by the CS player and O_i player, respectively.
$\hat{O}_{s,i}$	The preferred offer of the CS player to an SP_i player.
$\hat{O}_s$	The preferred offer of the CS player that is directed to the SP player that is best for the CS player.
$\hat{O}^-_{s,i}$	The SP_i preferred offer at state s.
P^*_s	A preferred (optimal) path for the CS player at a state s.
s^*	The state resulting in using the preferred path at state s ($Res(s, P^*_s)$).
pr	The probability that the CS player will succeed to reach the goal G_i.
$Eu_l(s, O)$	The expected utility of the $l \in \{CS, i, j\}$ agent from implementing an offer O at state s.
$Eu_l(s, P)$	The expected utility of the $l \in \{CS, i, j\}$ agent from the CS agent following path P at state s.
$sw_{i,cs}(s, O)$	The social welfare of an offer O at state s between the customer player and the SP_i player.

4. Game formalization

In this section, we provide a formalization of the contract game as well as the necessary definitions to characterize the equilibrium strategies of the contract game. A summary of all the notations appears in Table 1.

Given a board in the Contract Game, a location (x_1, y_1) is said to be *near* another location (x_2, y_2) if it is one move away, that is $|x_2 - x_1| + |y_2 - y_1| = 1$.

A state s of the game is a tuple: $\langle C_{cs}, C_1, C_2, (x, y), r, d\rangle$ where C_{cs} is the set of chips of the customer player, C_1 and C_2 are the sets of chips of the service provider players SP_g and SP_y respectively, (x, y) is the location of CS on the board, r is the round index in the game, and d indicates if the CS player moved in the previous round ($d = 0$ if it moved in the previous round, $d = 1$ indicates that it did not move in the previous round and $d = 2$ indicates it did not move for two consecutive round). There are two goal locations on the board: $G_1 = (\overline{x_1}, \overline{y_1})$ and $G_2 = (\overline{x_2}, \overline{y_2})$.

A *path* P from (x_1, y_1) to (x_k, y_k) on the board is defined as a sequence of locations on the board $\{(x_1, y_1), \ldots, (x_l, y_l), \ldots, (x_k, y_k)\}$ such that (x_l, y_l) is near (x_{l+1}, y_{l+1}) for any $1 \le l \le k-1$. As an example, in Fig. 1, we can see a possible path outlined on the board from the current location of the CS player to the SP_y service provider. We say a path P' is *contained* in a possible path P, denoted by $P' \subseteq P$ if all coordinates $\{(x_1, y_1), \ldots, (x_k, y_k)\}$ in P' are also coordinates in P.

The set of *needed chips* to go through a path P is denoted by C_P. A path P is *possible* in state s if $C_P \subseteq C_{cs}$ and $(x_1, y_1) = (x, y)$. Let $s = \langle C_{cs}, C_1, C_2, (x, y), r, d\rangle$ be a state and $P = \{(x, y), \ldots, (x_l, y_l), \ldots, (x_k, y_k)\}$ be a possible path of s. The *result* of CS moving according to P, denoted by $Res(s, P)$, is the state $\langle C_{cs} \setminus C_P, C_1, C_2, (x_k, y_k), r+1, d=0\rangle$. To illustrate, in Fig. 1, if the CS moves on the outlined path, this will result in the new state in which the location of the player will be at the goal square at the end of the path and the player has reduced from its chip set 9 gray chips.

An offer O is a pair (O_{cs}, O_i) $i \in \{1, 2\}$ such that $O_{cs} \subseteq C_{cs}$ is the set of chips that the customer CS will send to player SP_i and $O_i \subseteq C_i$ is the set of chips that the SP_i will send to the CS player. We extend the $Res(s, O)$ notation to account for an offer $O = (O_{cs}, O_i)$ that is accepted in state $s = \langle C_{cs}, C_1, C_2, (x, y), r, d\rangle$. For the case in which the proposal involves the SP_g player, then we have

$$Res(s, O) = \langle C_{cs} \cup O_i \setminus O_{cs}, C_1 \cup O_{cs} \setminus O_1, C_2, (x, y), r, d\rangle$$

and similarly for the case in which the proposal involves the SP_y player. For example, suppose a proposal from SP_g to the CS player at the state that is depicted in the board of Fig. 1 is to send 33 red chips and 7 purple chips for 11 gray chips. In this case, the resulting state from the acceptance of the offer includes 33 red chips, 12 yellow chips, 7 purple and 1 gray chip for the CS player, and 11 gray chips for the SP_y player (the SP player that is located on the yellow goal square in the figure).

Without loss of generality, throughout the paper we will refer to the yellow goal square as G_2 and the gray goal square as G_1. Similarly, SP_g refers to the service provider player who is located on the G_1 goal square and SP_y refers to the service provider who is located on the G_2 goal square.

The game ends in a terminal state $s = \langle C_{cs}, C_1, C_2, (x, y), r, d\rangle$ in which one of the following holds:

- The CS agent reached the G_1 goal.
- The CS agent reached the G_2 goal.
- The CS player has not moved for two consecutive rounds ($d = 2$) i.e., in the two states prior to s, the location of the CS was also (x, y).

A player's performance in the game is measured by a scoring function. Each player obtains c points for each chip it has at the end of the game. If the CS player reached one of the goals G_i then both the CS player and the service provider SP_i receive a bonus b^*. In the specific game that we played c was 5 and b^* was 150 points. For a terminal state s we denote by $u_i(s)$ the score of any player i at s. We extend u_i to non-terminal states to be $c \cdot |C_i|$ where $|C_i|$ represents the number of chips player i has.

Importantly, although reaching one of the goals of the service providers before the game terminates increases the score of the customer player, we do not assume that the customer player will necessarily follow this strategy. This allows us to capture players that exhibit boundedly rational behavior. In the next section, we shall see how this allows to define an equilibrium strategy that reasons about people's play. We denote by pr the probability that the CS player will succeed to reach a goal G_i following a possible path P toward G_i and receive the bonus b^*.

Let $Eu_l(s, O)$ indicate the expected utility of the $l \in \{CS, i, j\}$ agent from implementing an offer O at state s given that all players follow the subgame perfect equilibrium in state $Res(s, O)$. This equilibrium will be formalized in Theorem 1. Similarly, let $Eu_l(s, P)$ denote the expected utility of the $l \in \{CS, i, j\}$ agent from the CS agent following path P at state s given that all players follow the subgame perfect equilibrium in state $Res(s, P)$.

We will first define the expected utility of the players from the CS player attempting to reach the goal G_i by following the possible path P, given that the probability of getting to the goal is pr. The expected utility to SP_i is defined as:

$$Eu_i(s, P) = p \cdot u_i(Res(s, P)) + (1 - p) \cdot u_i(s) \tag{1}$$

For the SP_j agent ($j \neq i$) the expected utility is defined as follows, because the CS player will not reach its goal.

$$Eu_j(s, P) = u_j(s) \tag{2}$$

The expected utility of the CS player is:

$$Eu_{cs}(s, P) = p \cdot u_{cs}(Res(s, P)) + (1 - p) \cdot u_i(s) \tag{3}$$

We define the social welfare of an offer O at state s between a CS player and an SP_i player to be

$$sw_{i,cs}(s, O) = Eu_{cs}(s, O) + Eu_i(s, O) \tag{4}$$

To illustrate, suppose that in the state presented in Fig. 1, the CS player attempts to move toward the goal of the gray player G_1 using the path P marked with the black line. Suppose further that $p = 0.9$. If the CS player succeeds in its move than it is left with 3 gray chips and 12 yellow chips. The agent also gets the a bonus of 150 points. If the agent fails, it keeps all its 24 chips. Thus, its expected score is $0.9 \cdot (15 \cdot 5 + 150) + 0.1 \cdot 24 \cdot 5 = 214.5$. The expected score of SP_g is $0.9 \cdot (40 \cdot 5 + 150) + 0.1 \cdot 40 \cdot 5 = 335$.

4.1. Possible offers

In this section we define the set of offers that are possible for the CS player and an SP_i player, denoted *possible offers*. This definition depends on the notion of a *preferred path*, which due to the recursive nature of our equations we will formalize later, in Definition 5. A preferred path at state s, denoted P_s^*, is a possible path at s such that for any other possible path P at s, i.e., $Eu_{cs}(s, P) \leq Eu_{cs}(s, P_s^*)$. A possible offer O at state s between the CS player and an SP_i player will not make the CS worse off than following his preferred path P_s^* at state s. In a similar way, the possible offer O will not make the SP_i player worse off than the CS player following its preferred path.

Definition 1. An offer O is a possible offer for a CS and a given SP_i player at state s if the following conditions hold:

$$Eu_i(s, O) \geq Eu_i(s, P_s^*) \text{ and } Eu_{cs}(s, O) \geq Eu_{cs}(s, P_s^*) \tag{5}$$

The preferred offers for the CS player is a subset of possible offers that are stated in the following definition.

Definition 2 *(CS preferred offer toward SP_i).* An offer O involving the CS player and the SP_i player at state s is *a CS Preferred Offer to SP_i*, denoted $\hat{O}_{s,i}$, if the following conditions hold

1. $Eu_i(s, O) > Eu_i(s, P_s^*)$ and $Eu_{cs}(s, O) \geq Eu_{cs}(s, P_s^*)$.
2. For any offer O' toward SP_i that satisfies condition (1) it holds that $Eu_{cs}(s, O') \leq Eu_{cs}(s, O)$.

The preferred offer $\hat{O}_{s,i}$ is defined to be the empty set if no such offers O exist. Otherwise, $\hat{O}_{s,i}$ is defined as

$$\hat{O}_{s,i} = \arg\max_{O}\{Eu_i(s,O) \mid O \text{ is a } CS \text{ preferred offer to } SP_i \text{ at } s\} \tag{6}$$

Condition 1 of Definition 2 requires that O is a possible offer according to Definition 1. Condition 2 of Definition 2 requires that there is no alternative possible offer that is better for the CS player.

We extend this definition to describe the preferred offer for the CS player that is directed to the SP player that will provide the CS player with higher utility.

$$\hat{O}_s = \text{argmax}\{Eu_{cs}(s,\hat{O}_{s,g}), Eu_{cs}(s,\hat{O}_{s,y})\}. \tag{7}$$

We break ties by preferring offers involving an SP player that yield higher sum of utilities

$$sw_{i,cs}(s,\hat{O}) > sw_{j,cs}(s,\hat{O}) \tag{8}$$

If the sum of utilities is equal, then we prefer offers made to SP_g over offers made to SP_y. Note that ties can also be broken towards SP_y and the equilibrium definition will still hold. We expand on this point in the Discussion (Section 6.5).

The preferred offers for an SP_i player is a subset of possible offers that are strictly more beneficial to the CS player following the preferred path P^*_s and most beneficial for the SP_i. This is stated in the following definition.

Definition 3 *(SP_i preferred offer).* The SP_i and SP_j, $i \neq j$ preferred offers at state s, $\hat{O}^-_{s,i}$, $\hat{O}^-_{s,j}$ satisfy the following conditions:

1. $Eu_{cs}(s,\hat{O}^-_{s,i}) \geq Eu_{cs}(s,P^*_s)$, $Eu_i(s,\hat{O}^-_{s,i}) > u_i(s)$ and $Eu_i(s,\hat{O}^-_{s,i}) \geq Eu_i(s,\hat{O}^-_{s,j})$. If there is no offer that satisfies this condition we set $\hat{O}^-_{s,i} = (\emptyset,\emptyset)$.
2. $Eu_{cs}(s,\hat{O}^-_{s,j}) \geq Eu_{cs}(s,P^*_s)$, $Eu_j(s,\hat{O}^-_{s,j}) > u_j(s)$ and $Eu_j(s,\hat{O}^-_{s,j}) \geq Eu_j(s,\hat{O}^-_{s,i})$. If there is no offer that satisfies this condition we set $\hat{O}^-_{s,j} = (\emptyset,\emptyset)$.
3. For any offer O involving SP_i that satisfies the constraints in condition (1), if $Eu_i(s,O) > Eu_i(s,\hat{O}^-_{s,i})$ then either $Eu_{cs}(s,O) < Eu_{cs}(s,\hat{O}^-_{s,j})$, or $Eu_{cs}(s,\hat{O}^-_{s,j}) = Eu_{cs}(s,O)$ and ties were broken in favor of SP_j.
4. For any offer O involving SP_j that satisfies the constraints in condition (2), if $Eu_j(s,O) > Eu_j(s,\hat{O}^-_{s,j})$ then either $Eu_{cs}(s,O) < Eu_{cs}(s,\hat{O}^-_{s,i})$, or $Eu_{cs}(s,\hat{O}^-_{s,i}) = Eu_{cs}(s,O))$ and ties were broken in favor of SP_i.

Condition 1 of Definition 3 states that the benefit at state s from the SP_i preferred offer $\hat{O}^-_{s,i}$ is at least as high the benefit that is associated at state s for the SP_i player. In addition, for the SP_i player, its benefit from $\hat{O}^-_{s,i}$ must be at least as high as its benefit from the SP_j preferred offer $\hat{O}^-_{s,i}$. Condition 2 states symmetrical requirements regarding the benefit at state s from the SP_j preferred offer $\hat{O}^-_{s,j}$. Condition 3 states that if an alternative offer O is more beneficial to SP_i player than the SP_i preferred offer, then it must be the case that the CS player prefers the SP_j offer. Condition 4 sates symmetrical requirements regarding an alternative offer O for the SP_j player.

We now define the SP preferred that will provide the CS player with higher utility.

$$\hat{O}^-_s = \text{argmax}\{Eu_{cs}(s,\hat{O}^-_{s,g}), Eu_{cs}(s,\hat{O}^-_{s,y})\}. \tag{9}$$

We break ties by preferring offers that yield higher sum of utilities

$$sw_{i,cs}(s,\hat{O}^-) > sw_{j,cs}(s,\hat{O}^-) \tag{10}$$

If the sum of utilities is equal, then we prefer offers made by SP_g over offers made to SP_y.[1]

A special type of possible offers involve a commitment between one of the SP players and the CS player. Such offers prevent the CS player from reaching the goal square of the *other* SP player. The CS player is *committed* to player SP_i, $i \in \{1,2\}$ in state $s = \langle C_{cs}, C_1, C_2, (x,y), r, d\rangle$ if for any path P from (x,y) to G_j, $j \in \{1,2\}$, $j \neq i$, $C_P \not\subseteq (C_{cs} \cup C_j)$.

Definition 4 *(Commitment offer).* An offer $O = (O_{cs}, O_i)$ made at a state s is a *commitment offer* toward SP_i if the following conditions hold:

1. in s the CS player is not committed toward any of the SP players,
2. the resulting state $Res(s,O)$ is a commitment state towards SP_i.

[1] As stated above, ties can also be broken towards toe SP_y player.

Intuitively, if the CS player is committed to player SP_i in state s, then there is no possible path for the CS player at this state to the goal of the other SP_j player, even if the SP_j player provides all its chips to the CS player. Therefore the only way for the CS player to receive the bonus in the game is to reach the goal of the SP_i player. As an example, a commitment offer at the beginning of the game shown in Fig. 1 is when the SP_g proposes to send 33 red chips and 7 purple chips for 11 yellow chips. As we will show, commitment offers play a key role in the equilibria strategies that are described in the following section.

4.2. Preferred paths

We are now ready to formalize the preferred path P_s^* for any state s. We will use the following notation. Let $s^* = Res(s, P_s^*)$ denote the state resulting in the CS using the preferred path at state s.

Definition 5 *(Preferred path).* We define the preferred path at state s as the best of four possible candidate paths $\{P_1, P_2, P_3, P_4\}$ that are defined as follows.

1. If the CS player is able to reach one of the goals at state s, then P_1 is the shortest path towards the goal square, breaking ties in favor of the path leading to SP_y. Otherwise $P_1 = \emptyset$.
2. For any path P that does not lead to one of the goals (and if the number of dormant turns d is 1, then $P \neq \emptyset$) one of the following must hold:
   (a) if it is the SP's turn to make offers in $Res(s, P)$ then:
   $$P_2 = \operatorname*{argmax}_P \{Eu_{cs}(Res(s, P), \hat{O}^-_{Res(s,P)}))\}$$
   Let $EU_l(s, P_2) = Eu_l(Res(s, P_2), \hat{O}^-_{Res(s,P_2)})$ denote the expected utility to any player from the CS player choosing path P_2.
   (b) if it is the CS' turn to make offers in $Res(s, P)$ then:
   $$P_3 = \operatorname*{argmax}_P \{Eu_{cs}(Res(s, P), \hat{O}_{Res(s,P)})\}$$
   Let $EU_l(s, P_3) = Eu_l(Res(s, P_3), \hat{O}_{Res(s,P_3)})$
   We break ties toward shorter paths.
3. If the CS player must move in s, i.e., $d = 1$, $P_4 = \emptyset$ and $Eu_l(P_4) = u_l(s)$.

We choose $P_s^* = \operatorname{argmax}_{P \in \{P_1,..P_4\}}\{Eu_{cs}(s, P)\}$. We break ties in favor of paths that leads to one of the goals, and otherwise in favor of short paths.

Using the expected utility of preferred paths we can specify the expected utility of implementing an offer O in state s for player $l \in \{cs, 1, 2\}$ as follows.

$$Eu_l(s, O) = Eu_l(Res(s, O), P^*_{Res(s,O)}) \tag{11}$$

The main question is whether P_s^* is well defined since the definition is recursive and we need to show that there are no circles in the definition. This is stated in the next proposition.

Proposition 1. *For any state s the preferred path P_s^* is well defined.*

The proof of this proposition is as follows. First, the contract game always ends after a finite number of rounds. Since the number of chips of the players is finite, the upper bound on the number of rounds is twice the sum of the number of chips of all players. Furthermore, the same state cannot appear twice in a game. Either, the SP moved and the total number of chips decreased by one. If it doesn't move, then d is increased which leads to a different state.

4.3. Examples

We will demonstrate the preferred paths and preferred offers using a few examples. In all examples we assume that the CS player always succeeds to reach the goal (i.e., $p = 1$). Consider the board in Fig. 2 and suppose that the state s corresponds to each of the players having 10 red chips, and that the SP_y player (whose goal is $G2$) has an additional 1 yellow chip, and the SP_g player (whose goal is $G1$) has an additional 1 gray chip. Suppose the CS player (the "me" player) must move (i.e., the number of dormant turns is $d = 1$) and it is its turn to make an offer. The CS preferred offer (Definition 2) depends on its preferred path (Definition 5). Since it does not have the needed chips to move, the preferred path of the CS player in state s is the empty set, and its expected utility $Eu_{cs}(s, P_s^*)$ is defined in Equation (4) as $5 \cdot 10 = 50$. Similarly, the expected utility for each of the SP players $Eu_i(s, P_s^*)$ is defined in Equation (4) as $5 \cdot 11 = 55$. The CS preferred

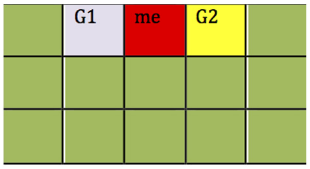


Fig. 2. Example demonstrating preferred offers and preferred paths. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

offer to the SP_g requests all of the chips (i.e., one gray and 10 red chips) from the SP_g player. Similarly, the CS preferred offer to the SP_y requests all of the chips from the SP_y player. Ties are broken towards the SP_g player, since the sum of the utilities for the CS preferred offer to both SP_g and SP_y players is equal (Equation (8)). The SP_g will accept the offer and the CS player will move one square to the gray goal, thereby ending the game. The final score for the CS player will be $20 \cdot 5 + 150 = 250$ points, the final score for the SP_g player will be 150 points, and the final score for the SP_y player will be 55 points.

If it is the turn of the SP players to make an offer, then the SP_g player will make the SP_g preferred offer $\hat{O}^-_{s,g}$ of Definition 3, which is to offer all of its chips to the CS player. The SP_y player will make the SP_y preferred offer $\hat{O}^-_{s,y}$ of Definition 3, which is to offer all of its chips to the CS player. The score that is associated with accepting the SP_g preferred offer for the CS player is 250 points, which is strictly higher than the utility at state s, which is 50 points. The utility to the SP_g player from this offer is 150 points, which is strictly greater than its utility at state s, which is 55 points. In addition, the utility to the SP_g player from this offer is greater than its utility given that the CS player accepts the SP_y preferred offer, which is also 55 points. Hence Condition 1 of Definition 3 is satisfied. Any offer that is better off for the SP_g player and satisfies Condition 1 of Definition 3 (e.g., offering only 9 chips to the CS player) is worse off for the CS player than accepting the SP_y preferred offer. Hence Condition 3 of Definition 3 is satisfied. In a similar way, the SP_y and SP_g preferred offers satisfy Conditions 2 and 4 of Definition 3. The CS player, breaking ties towards SP_g, will accept the offer of the SP_g player.

Suppose now that the board is the one in Fig. 2, but that the chip distribution is as follows: all players have 10 red chips, and in addition the CS player has 1 yellow chip and 1 gray chip. Suppose again that it is the turn of the CS player to make an offer and that it must move ($d = 1$). In the current state there are two viable paths for the CS player, one leading to the goal of the SP_g player and one leading to the goal of the SP_y player. The preferred path of the CS player breaks ties toward the SP_y player and the score of the SP_y player will be $5 \cdot 10 + 150 = 200$, while the score for the SP_g player will be $5 \cdot 10 = 50$. The CS preferred offer towards the SP_g player will offer one yellow chip for all of its 10 red chips. This offer represents a commitment (Definition 4) of the CS player towards the SP_g player. The score for the CS player for this offer will be $20 \cdot 5 + 150 = 250$ and for the SP_g player will be $1 \cdot 5 + 150 = 155$. Note that sending its yellow chip is a necessary condition for the commitment to hold. If this offer did not include a yellow chip, then SP_g should not accept the offer, because the preferred path (and the only viable path towards the goal) for the CS player would lead it to the goal of the SP_y player (because ties are broken towards the SP_y player).

Lastly, suppose that the board is the one in Fig. 2, but the SP_g player has 1 gray chip and all players have 10 red chips. Suppose that it is the turn of the SP players to make an offer. In this case, the SP_g preferred offer will provide one gray chip in return for all of the 10 red chips. The score of the SP_g player will be $20 \cdot 5 + 150 = 250$ and the score of the CS player will be 150. If the CS player rejects the offer then its score will be 50; therefore the CS player is strictly better off accepting the offer. The SP_y player cannot make an offer that will allow the CS player to reach the goal.

5. Equilibrium strategies

In this section we present an equilibrium definition for the contract game. We first provide the strategies for any possible contract game. Then, we will present the details of the strategies for the specific game we ran our experiment on.

5.1. Sub-game perfect equilibrium

We first provide a general description of the equilibrium strategies for both player types. The strategy depends on the features mentioned in the beginning of section 4.1.

Theorem 1. *The following strategies form a sub-game perfect equilibrium for the contract game.*

Strategy for the SP_i player: *Given a state $s = \langle C_{cs}, C_1, C_2, (x, y), r \rangle$, the strategy for the SP_i is as follows:*

1. *If the SP_i player receives the offer O_i from the CS player, it will accept it only if the following holds: $Eu_i(s, O_i) > Eu_i(s, P^*_s)$.*
2. *If it is the turn of the SP_i player to make the offer, it will make the SP_i preferred offer $\hat{O}^-_{s,i}$.*

Strategy for the CS player: *Given a state $s = \langle C_{cs}, C_1, C_2, (x, y), r\rangle$, the strategy for the CS is as follows:*

1. *If it is the negotiation stage of an odd round and it has received offers O_2 and O_1 and let O_l denote the best offer for the CS player (breaking ties first towards the offer that maximizes the social welfare (Equation (8)) and then towards the offer that is made by SP_g player):*

$$O_l = \arg\max_{O_i \in \{O_2, O_1\}} Eu_{cs}(s, O_i) \tag{12}$$

The CS player will accept O_l only if

$$Eu_{cs}(s, O_l) \geq Eu_{cs}(s, P_s^*). \tag{13}$$

2. *If it is a negotiation stage of an even round (the CS makes the proposal) then make the proposal $\hat{O}_s$ toward the relevant SP_i.*
3. **Movement:**
If it is a movement state then move according to the preferred path P_s^*.

Note that ties are broken towards the SP_g player when making and responding to offers, but are broken towards the SP_y player when choosing paths on the board. The reason for this was to provide each service provider agent with fair opportunity for an advantage in the game, rather than consistently favoring one of the service providers at all decision points. From a theoretical point of view we can change the breaking ties arbitrarily but keep all the definitions consistent. The correctness of Theorem 1 will not be affected. For the remainder of this section, we use the term "equilibrium strategy" to refer to the one that is defined in Theorem 1.

5.2. Proof of Theorem 1

The correctness of Theorem 1 depends on the definition of the expected utility to the CS and the SP players. The following propositions states that the definitions of these utility functions are well defined.

Proposition 2 *(Correctness of expected utility definitions). Let O be an offer that is accepted at state s, P_s^* is a preferred path that is followed by the CS player starting at state s and that both CS and SP players follow the equilibrium strategies defined in Theorem 1. Then the expected utilities $Eu_i(s, P)$, $Eu_j(s, P)$, $Eu_{cs}(s, P)$, $Eu_i(s, O)$, $Eu_j(s, O)$ are correctly defined by Equations (1), (3), (2), (11) and Definition 5.*

The proof follows from the equilibrium strategy defined in Theorem 1 and from the definitions of the aforementioned equations.

We now show that the SP does not have an incentive to deviate from its equilibrium strategy in Theorem 1.

Strategy of the SP player for receiving an offer: If the SP_i will reject the offer O_i, the game will proceed to the movement phase. According to the equilibrium strategy for the CS player, it will follow P_s^*. According to Proposition 2, the expected utility to the SP_i player from accepting the offer O_i is $Eu_i(s, O_i)$ and his expected utility from the CS player following the path P_s^* is $Eu_i(s, P_s^*)$. According to condition 1 in the equilibrium definition of the strategy for the SP_i player, we have that $Eu_i(s, O_i) > Eu_i(s, P_s^*)$, and thus deviation by rejecting the offer is not beneficial.

Strategy for the SP to make an offer: In this case SP_i offers $\hat{O}^-_{s,i}$ and SP_j offers $\hat{O}^-_{s,j}$. Deviation by SP_i and offering another offer O is beneficial only if $Eu_i(s, O) > Eu_i(s, \hat{O}^-_{s,i})$ and either $Eu_{cs}(s, O) > Eu_{cs}(s, \hat{O}^-_{s,j})$ or if $Eu_{cs}(s, \hat{O}^-_{s,j}) = Eu_{cs}(s, O))$ ties are broken in favor of SP_i. Otherwise, the offer will not be accepted or won't be beneficial. According to Definition 3 such O exists only if it does not satisfy the conditions $Eu_{cs}(s, O) \geq Eu_{cs}(s, P_s^*)$, $Eu_i(s, O) > u_i(s)$ and $Eu_i(s, O) \geq Eu_i(s, \hat{O}^-_{s,j})$. First, if $Eu_{cs}(s, O) < Eu_{cs}(s, P_s^*)$ then the offer will not be accepted. So, either the score for SP_i will be $u_i(s)$, but $Eu_i(s, \hat{O}^-_{s,i}) > u_i(s)$ or the expected utility will be $Eu_i(s, \hat{O}^-_{s,j})$ but $Eu_i(s, \hat{O}^-_{s,i}) \geq Eu_i(s, \hat{O}^-_{s,j})$. Thus, in both cases deviation is not beneficial. Clearly, if $Eu_i(s, O) \leq u_i(s)$ or $Eu_i(s, O) < Eu_i(s, \hat{O}^-_{s,j})$ deviation is not beneficial since $Eu_i(s, \hat{O}^-_{s,i}) > u_i(s)$ and $Eu_i(s, \hat{O}^-_{s,i}) \geq Eu_i(s, \hat{O}^-_{s,j})$.

Strategy for the CS player to receive an offer: If the CS player rejects the offer O_l, the game will proceed to the movement phase. According to the equilibrium strategy for the CS player it will follow P_s^*. According to Proposition 2, the expected utility to the CS player from accepting the offer O_l is $Eu_{cs}(s, O_l)$ and his expected utility from following the path P_s^* is $Eu_{cs}(s, P_s^*)$. According to condition 13 in the equilibrium definition of the strategy for the SP_i player, we have that $Eu_{cs}(s, O_i) \geq Eu_{cs}(s, P_s^*)$. Thus deviation by rejecting the offer is not beneficial.

Strategy for the CS player to make an offer: Suppose that the CS's player offer $\hat{O}_s$ is toward SP_i. By the definition $Eu_i(s, \hat{O}_s) > Eu_i(s, P_s^*)$ and this offer will be accepted by SP_i. If CS player will make an offer that will not be accepted by the SP_i then its expected utility will be $Eu_i(s, P_s^*)$ and again by the definitions $Eu_{cs}(s, \hat{O}_s) \geq Eu_{cs}(s, P_s^*)$. So, beneficial deviation O' given to SP_i should satisfy $Eu_i(s, O') > Eu_i(s, P_s^*)$, $Eu_i(s, O') > Eu_{cs}(s, \hat{O}_s)$ but according to the preferred offer Definition 2 such an offer does not exist. Furthermore, according to equation (7) such an offer toward SP_j does not exist either.

Strategy for the CS player to move: We now show that it is preferable for the CS player to move using P_s^*. The CS player can either go directly toward the goal, if possible or stay in its current location or move according to one of the possible paths that are not toward the goal. These are the cases that are considered in Definition 5 and P_s^* is defined as the one that maximizes the expected utility from all these options. Furthermore, according to Proposition 2, the expected utility of P_s^* is correct, deviation is not beneficial.

6. Empirical study

In this section we describe the evaluation of several types of agents for playing the contract game with human players. We recruited 500 human subjects from three countries: Israel, the U.S. and China. These included 208 computer science or engineering undergraduates from two Israeli universities (average age was 25 years; 35% female), 149 students from the greater Boston area (average age was 22; 48% female), and 143 students from universities in Beijing, China (average age was 23; 46% female).

Each participant was given an identical 25-minute tutorial on the contract game as well as a 5-minute quiz about the rules of the game.[2] We used identical versions of the tutorial in three languages, Hebrew, English and Mandarin. Participants were seated in front of terminals for the duration of the study, and could not speak to each other or see each other's terminals. The tutorial consisted of a detailed explanation of the rules of the game in the form of a slide presentation that participants could go through at their own pace. By design, It did not include any examples of negotiation strategies to avoid priming subjects. To ensure that participants understand the game, we included a quiz at the end of the tutorial, and participating in the study was contingent on passing the quiz. To avoid long games, there was a 5-minute timeout for interacting during the communication phase in each round. At this point the game automatically proceeded to the movement phase (which was also associated with a timeout). Participants could see a counter representing the amount of time left to play in each of the phases.

All participants were told they would be playing the game with other people or with other computer agents. This wording was chosen to standardize the instructions for all players. It reflects situations as in e-commerce in which players may not know the identity of their negotiation partners. Players were paid a fixed sum of 12 U.S. Dollars (in local currency) for participating in the game. We did not pay people with respect to their performance for two reasons: First, our goal was to create an agent that can play well with people and not to claim that their behavior was completely driven by their preferences. Second, past work shows that people play competitively in CT games when given a fixed reward and their play does not differ substantially from participants whose reward depends on their score [11].

We ran two types of configurations, one consisting of all human players and the other consisting of two people and a computer agent playing service provider or customer role. All of the games played included three human players, or a single human player and two computer agents playing the various roles. Specifically, we played games consisting of 3-human players (denoted "HvsH"), games consisting of an agent playing the customer role and two human players (denoted "HvsCSa"), games consisting of an agent playing the service provider role and two human players (denoted as "HvsSPa"). In all games involving a human and a computer agent playing the service provider role (i.e., HvsSP-A, Hvs-RAP, Hvs-RAPD), the computer agent was consistently assigned to play the role of the SP_y player. Note that this is without loss of generality, as the SP_g and SP_y players are symmetrical in the dependency relationship with the CS player. They both have the same chips and are located at equal distance from the position of the CS player.

6.1. Board game and commitments

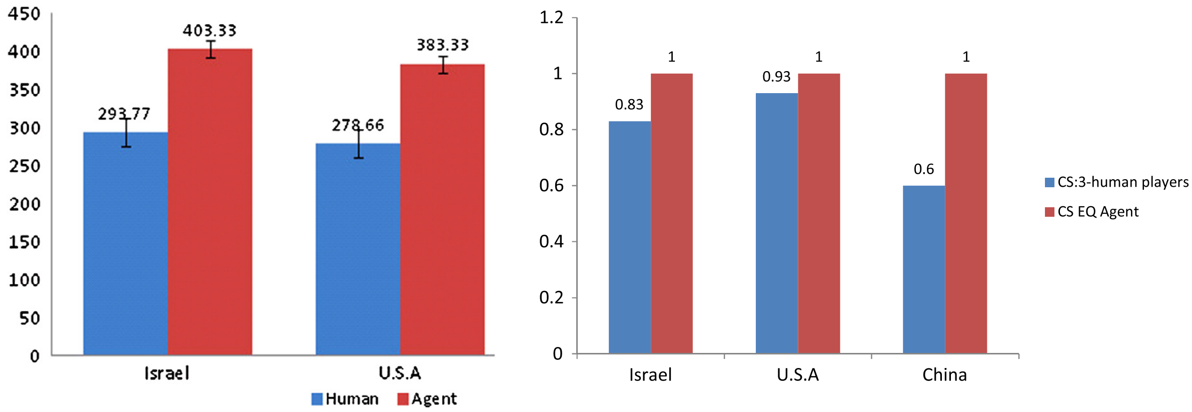
We used the board game of Fig. 1 for all of the studies. In this board, the CS is independent (has the resources to get to either of the goals at the onset of the game), while the SP players are symmetric (they have the same number of chips and the length of the path to each of their goals is equal). In this game, the preferred offers for the SP and the CS player are commitment offers. Furthermore, according to the sub-game equilibrium of Theorem 1, both SP players offer their preferred offer, the CS player accepts the offer of the SP_g player, and proceeds directly to the goal, thereby ending the game.

For the customer (CS) role, we used an agent termed "CS-A" that used the strategy that is described in the equilibrium definition of Theorem 1 while assuming the customer player always reaches the goal (i.e., $p = 1$) after a commitment has been made. For the service provider (SP) role, we used one of two types of equilibrium agents. One agent, termed "SP-A", used the strategy that is described in the equilibrium definition in Theorem 1, while assigning probability $p = 1$ for the customer player reaching the goal.

The agent termed "SP-RAP" extended the SP-A agent in two ways in order to handle the uncertainty that characterizes human play in negotiation settings. First, it used a risk averse strategy using a concave utility function $1 - e^{-\alpha \cdot u_i(s)}$ where $u_i(s)$ is the score of the SP_i player at state s. Second, it reasoned about a possibly bounded rational CS player by assigning a positive probability $p > 0$ for the customer player not reaching the goal. A separate value for pr was assigned for each country by dividing the number of times the CS player reached the goal by the total number of games played.[3] We measured

[2] The tutorial and quiz can be downloaded at the link http://tinyurl.com/zp4lapk.

[3] In all countries, there were some cases in which the customer player did not reach the goal despite having the required resources at the onset of the game.



(a) Average Score (b) Reaching Goal

Fig. 3. Performance of the CS-A agent.

the parameter *pr* on a set of held out games which included only humans. The lowest parameter setting was obtained in China, where *pr* was set at 0.6. In Israel, the value of the *pr* parameter was set at 0.83. The value obtained for *pr* in the USA was significantly higher than that of China and Israel and was set at 0.93.

We highlight the difference in the strategy of these different agents by specifying the first preferred commitment. The preferred commitment offer of the SP-RAP agent depended on the value of the *pr* parameter and thus the offer made the SP-RAP agent was different in each country. In China, the preferred commitment offer made by the SP-RAP agent offered 19 chips (12 red chips and 7 purple chips) in return for 11 gray chips. In Israel, the preferred commitment offer made by the SP-RAP agent offered 26 chips (19 red chips and 7 purple chips) for 11 gray chips. In the USA, the preferred commitment offer made by the SP-RAP agent offered 33 chips (26 red chips and 7 purple chips) for 11 gray chips. Our hypothesis was that the SP-A agent (which does not reason about bounded rational customer players) will be outperformed by the SP-RAP agent when playing with people in the customer role. The SP-A agent offered 40 chips (33 red chips and 7 purple chips) in return for 11 gray chips. The Appendix shows how to compute this preferred commitment offer for the SP-A agent. The process for computing the preferred commitment offers for the other agents is similar. All results in the sections that follow were statistically significant in the $p < 0.05$ range using t- and Mann Whitney (Wilcoxon) tests for unpaired data.

6.2. Analysis of performance for the customer role

We begin by comparing the performance of the CS-A agent to that of people playing with other people. Fig. 3a shows the average score of the CS-A agent in HvsCS-A games when compared to people playing the customer role in the HvsH games. As shown in the Figure, the CS-A agent outperformed people playing the customer role in all countries by a significant margin. In Israel, there was a significant difference in the score of the CS-A agent score (Mean = 403.33, SD = 48) and people (Mean = 293.77, SD = 20), $t(49) = 4.12$, $p < 0.0001$. In the U.S., there was a significant difference in the score of the CS-A agent (Mean = 383.33, SD = 72) and people (Mean = 278.6677, SD = 89), $t(28) = 3.53$, $p < 0.0001$. In China, there was a significant difference in the score of the CS-A agent (Mean = 364.76, SD = 72.4) and people (Mean = 230.66, SD = 106.43), $t(30) = 3.80$, $p < 0.0001$.

There are several possible factors that can explain the success of the CS-A agent, which we describe in turn. The first factor is the extent to which human players reach the goal and collect the bonus when playing the CS-A role. Fig. 3b compares the percentage of games in which the customer player reached the goal in HvsCS-A games and games in which human players played the customer role. As shown in the figure, the CS-A agent was always able to reach one of the service providers. This was not the case for the human players in the customer role. In particular, in China only 60% of human CS players reached the goal, followed by Israel (83%) and the U.S. (93%).[4] This result is striking, given that customer players have the necessary resources to reach the goal at the onset of the game, demonstrating irrational behavior by people during the game.

The second possible factor that may explain the success of the CS-A agent is the difference in benefit associated with the offers that were made by the equilibrium agents and those with the relevant human players. We define the *competitiveness measure* of an offer $O = (O_{cs}, O_i)$ made by the customer player to a service provider SP_i to equal to the difference between the number of chips sent by the SP_i and the number of chips sent by the customer player ($|O_i| - |O_{cs}|$). For example, suppose that the CS-A agent player proposes a commitment offer and asks for 40 red chips and proposes to send 11 yellow chips. In this case, its competitiveness measure will be 29 chips.

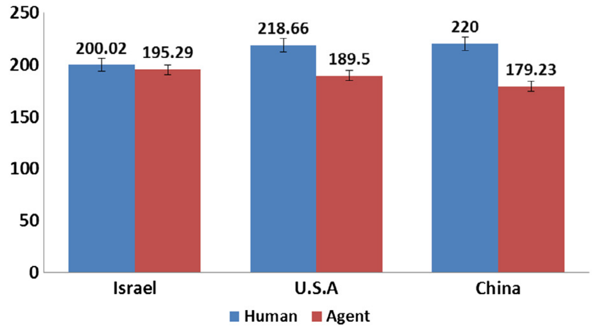
[4] We present a detailed comparison of people's behavior across countries in the next section.

Table 2
Average competitiveness of CS offers.

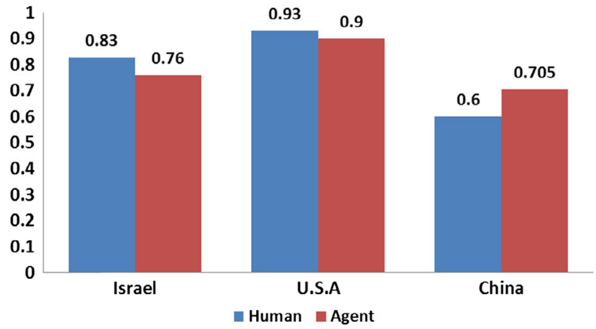
	HvsH	HvsCS-A	HvsSP-A	HvsSP-RAP
Competitiveness	6.67	**23.1**	9.15	6.76

Table 3
The percentage of games ended with a commitment agreement.

	HvsH	HvsCS-A	HvsSP-A	HvsSP-RAP
Games ended with commitments	31.8%	**68.7**%	73.6%	71.4%



(a) Average Score



(b) Reaching Goal

Fig. 4. Performance of the SP-A agent.

Table 2 compares the average competitive measure between offers made by the CS-A agent and offers made by people playing the customer role across in the different settings. The results are averaged over all countries. As shown by the table, offers made by the CS-A agent were significantly more competitive than offers made by people playing the same role.

To distinguish between the effects of getting to the goal and making competitive offers on the performance of the CS-A agent, we compared the final score in the game for the CS-A players. We controlled for those games in which the human CS player was able to reach the goal. The average score for the CS player in the HvsH setting was 315 points in Israel, 288 points in the U.S., and 294 points in China. Thus, in all countries, the average score for the CS-A agent (Fig. 3a) was consistently higher than that of people playing the same role that reached the goal. Therefore, we can conclude that the success of the CS-A agent cannot be solely attributed to the fact that people were not always able to reach the goal.

The third factor that explains the performance of the CS-A agent is how commitment offers were accepted by human players. According to the equilibrium strategy of Theorem 1 the customer player proceeds directly to the goal following an accepted commitment offer, effectively ending the game. Table 3 shows the ratio of games that ended following commitment offers was significantly higher in the HvsCS-A games (68.7%) than in the HvsH games (31.8%). As we have shown in Table 2, such commitment offers were significantly more beneficial to the CS-A agent than people playing the same role, and the CS-A agent proceeded to the goal square and ended the game after commitment offers were made. Together, these three factors provide a comprehensive explanation of how the CS-A agent was able to outperform people playing the same role.

6.3. Analysis of performance for service provider role

We now compare the performance of the service provider agent to people playing the same role. We begin with the SP-A agent which used the equilibrium strategy of Theorem 1 which assumed that the customer player always reaches the goal following the acceptance of a commitment offer (an assigned probability $pr = 1$).

Fig. 4a compares the performance of the SP-A agent to people playing the service provider role in all countries. In contrast to the performance of the CS-A agent, there was no statistically significant difference between the scores of the SP-A agent and people in all countries. The average score for the SP-A agent in each country was higher than that of people playing the service provider role.

To understand the reason for this poorer performance, we computed the average competitiveness of offers made by the SP-A agent. Because both service providers are competing to sign a contract with the customer player, we expected that offers made by the service provider role will be highly generous (i.e., will incur a negative competitiveness measure). As shown in Table 4, the SP-A agent was significantly less competitive than people playing the same role (-28 competitive ratio of the SP-A agent vs. -2.33 of people). Table 3 shows that the percentage of accepted commitments leading to game termination in HvsSP-A games (73.6%) was higher than in HvsCS-A games (65.9%). This means that the SP-A player made significantly more generous offers than people did, and that these offers were likely to be accepted by people. Lastly, as

Table 4
Average competitiveness of SP offers.

	HvsH games	HvsCS-A games	HvsSP-A games	HvsSP-RAP
Competitiveness	−2.33	−1.53	−28.78	−12.59

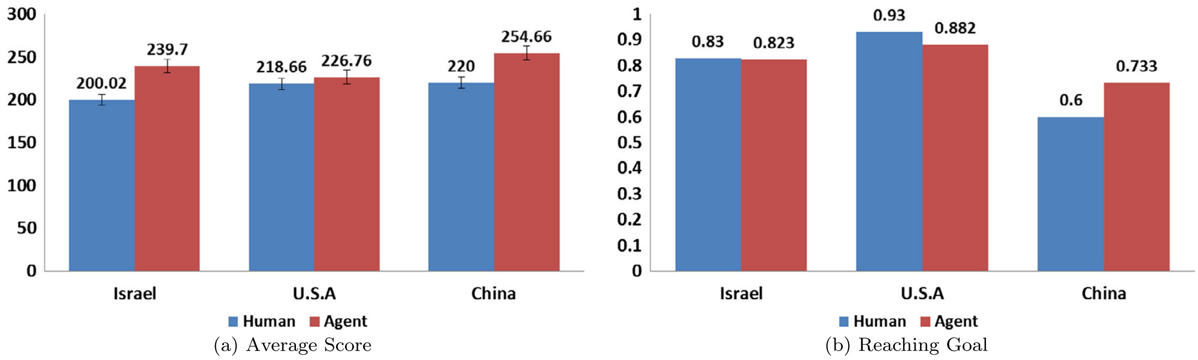


Fig. 5. Performance of the SP-RAP agent.

shown in Fig. 4b, human players playing the CS role in the HvSP-A games reached the goal significantly less than human players playing the CS role in the HvsH games. Together, these results explain the poor performance of the SP-A player.

Finally, Fig. 5a compares the performance of the SP-RAP agent to people playing the service provider role in all countries. As shown by the figure, the SP-RAP agent was able to outperform people in all countries. In Israel, there was a significant difference in the score of the SP-RAP agent score (Mean = 239.70, SD = 53.8) and people (Mean = 200.02, SD = 78.16), $p < 0.03$. In China, there was a significant difference in the score of the SP-RAP agent score (Mean = 254.66, SD = 65.9) and people (Mean = 220, SD = 59.3), $p = 0.05$. In the US the difference between the SP-RAP and the SP-A players was not statistically significant in the $p < 0.05$ range.

To explain the success of the SP-RAP, we note the following. First, the SP-RAP agent was significantly less generous than the SP-AP agent, while still more generous than people playing the same role. This can be seen in Table 4, which shows that the generosity measure of the SP-RAP agent was 14.5, compared to 28.78 for the SP-AP agent, and 2.33 for people. Second, as shown in Table 3, the ratio of accepted commitment offers in HvsSP-RAP games (71.4%) was significantly higher than those in HvsH games (31.8%) while not significantly different than those of HvsCS-A games (73.6%). Fig. 5b shows that human players in the CS role playing the SP-RAP agent were able to reach the goal as often (Israel and the U.S.) or more often than (China) people in the equivalent SP role in the HvsH games.[5] Together, these results show that reasoning about people not being able to reach the goal square resulted in less generous commitments by the SP-RAP player (compared to the SP-AP agent), but this did not affect the acceptance rate of these commitment proposals.

The next question to ask is whether we can improve the strategies of the CS-A and SP-RAP agents. While the SP-RAP agent's offers were generally less generous (i.e., more competitive) than those of the SP-A agent, they were less competitive than those of humans playing one of the SP roles (with associated benefit −2.33 in HvsH games vs −12.59 of the SP-RAP). Furthermore, even though the SP-RAP's offers in China were much less generous than those in Israel and the USA (with associated benefit of −14.5 in Israel; −15.56 in USA and −7.41 in China), the percentage number of games that ended with a human CS accepting the commitment was similar across al countries. Therefore we can attribute the higher score of the SP-RAP agent to the fact that it made less generous offers than the SP-A agent, yet they were still accepted by people. However, we hypothesized that as the SP-RAP agent's offers become less generous, at some point the human will prefer the offers of the other SP, which in turn reduces the score of the SP-RAP. Thus, it would be beneficial to quantify the extent to which the SP agent can be generous while still making offers that are accepted by human CS players, and are able to compete with the other service provider. Given the inherent lack of data in our domain (there was less than 50 offers made by the SP player in the HvsH setting in Israel) we cannot use classical machine learning algorithms for finding out the tradeoff.

Therefore we used a data-driven approach described as follows for the SP_i player (without loss of generality). At a state s, we seek an SP_i preferred offer denoted $\widehat{OT}_{s,i}$ with benefit $u_{cs}(s, \widehat{OT}_{s,i})$ to the CS player that represents a minimal acceptance threshold for the CS player, while still being likely to be accepted by people. To this end, we require that the following conditions are met: This first condition requires that the offer $\widehat{OT}_{s,i}$ is likely to be accepted by people. Let $\mathcal{O}_i$ be the set of offers between the SP_i player and the CS player made by people in the HvsH setting. The first condition

[5] There was no statistically significant difference in the percentage of people reaching the goal in the CS role between the HvsH and the HvsRAP setting both in Israel and the U.S. The percentage of people reaching the goal in the CS role in the HvsRAP setting in China was significantly higher than that of people playing the S role in the HvsH setting.

Table 5
Cultural differences in players' behavior in the game.

	Getting to goal	Commitment offers	Games ending with commitments
Israel	83%	26%	69%
USA	**90**%	19%	**73**%
China	68%	**28**%	67%

requires that any offer $O' \in \mathcal{O}_i$ at state s that provides equal or greater benefit to the CS player ($u_{cs}(s, O') \geq u_{cs}(s, \widehat{OT}_{s,i})$) was accepted by the CS player in the HvsH setting.

The second condition requires that the offer $\widehat{OT}_{s,i}$ is sufficiently generous to compete with the offers that can be made by the SP_j player in the HvsH setting. Recall that the generosity measure of an offer $O = (O_{cs}, O_i)$ made by an SP_i towards the CS player is the difference in the number of chips sent by the SP_i and the number of chips sent by the customer player ($|O_i| - |O_{cs}|$). This condition requires that the generosity measure of the offer $\widehat{OT}_{s,i}$ is at least as high as the average generosity of all $\mathcal{O}_j$ offers made by people playing the role of the SP_j player in the HvsH setting. This represents a baseline for offers to compete with those offers that are made by the other service providers.

We found that an offer $\widehat{OT}_{s,i}$ with a benefit threshold of 50 to the customer player (i.e., $u_{cs}(s, \widehat{OT}_{s,i}) = 50$) satisfied both of the conditions defined above. We therefore programmed a data-driven agent termed SP-RAPD that makes offers at state s in the game as follows: If a preferred offer exists at s that provides the offer $\widehat{OT}_{s,i}$ with a minimal acceptance threshold, then the agent makes the offer $\widehat{OT}_{s,i}$ (otherwise, it does not make any offer). In all other respects (accepting offers, moving on the board) the agent behaves like the SP-RAP agent. We illustrate the behavior of the SP-RAPD agent at the onset of the board game of Fig. 1. This agent would offer to the CS player the following commitment offer: 21 chips (14 red chips and 7 purple chips) in return for 11 gray chips. Note that this offer is less competitive than the commitment offer made by the SP-RAP agent in this state, which offered 26 chips to the CS player in return for the same 11 gray chips.

We compared the performance of this agent with that of people in Israel, playing 16 games in a lab setting that we will denote HvsRAPD. All subjects were engineering students from Ben-Gurion University from the same pool as the subjects in the other settings described earlier. The results from the games played in the HvsRAPD were as follows: As expected, the average generosity measure for the SP-RAPD agent was 10, which was lower than the average generosity measure for the SP-RAP agent that was reported in the HvsRAP games played in Israel (14.5). However, there was no difference in the percentage of accepted commitments made by the SP-RAPD agent (77%) and those made by the SP-RAP agent in Israel (76%). Consequently, the SP-RAPD agent was able to achieve an average score of 265.3 points (SD = 62.8), which was significantly higher than the average score achieved by the SP-RAP agent in Israel (239.7 points, SD = 53.8, $p < 0.05$) see Fig. 5a).

These results show that the SP-RAPD was able to adopt the right tradeoff between making beneficial offers to itself, while ensuring that the offers are accepted by people. In addition, the generosity measure of offers made by the SP_g human player was only 5.65, showing that the offers made by the SP-RAPD agent were still competitive with those offers made by the other service provider. Together, these results explain the success of the SP-RAPD agent in Israel. It demonstrates the potential of using data-driven heuristics to augment equilibrium agents in the absence of sufficient data.

6.4. Cultural differences in people's play

We end this section with studying cultural difference between people's behavior across all game conditions (HvsH, HvsSP-A and HvsSP-RAP games). To this end, we compare people's behavior in Israel, the USA and China using the following three measures: the extent to which people reached the goal when playing the role of the customer player, the ratio of commitment offers made by people versus the ones made by the agents and the extent to which people accepted commitment offers. The results are shown in Table 5.

As shown in the table, people in the US were most likely to reach the goal, followed by people in Israel and in China.[6] We observed two classes of scenarios in which people's irrational behavior prevented them from getting to the goal. The first consists of people who mistakenly send the needed chips to one of the SP players and are not able to reach the goal. The second scenario consists of human players who become too involved in the negotiations and end the game due to the dormant turn timeout.

It is also shown in the table that people made relatively more commitment offers in China, followed by Israel and then USA. Finally, people in the USA accepted more commitment offers that those in Israel and China, leading to the end of the game.

Finally, it is interesting to observe that the number of rounds in the games was not significant different between the countries. For example, in HvsH games, the average number of rounds in Israel was 5.86, in the China 5.33 and in USA it was 4.13.

[6] Note that the percentages here are slightly different from those used for pr in the SP-RAP strategy. The original ones were taken from only HvsH data, while Table 5 reports the results of all type of games where a human played the CS role.

6.5. Discussion

The fact that the CS-A agent who used sub-game perfect equilibrium was able to outperform people was striking. This contradicts past results in the literature showing that agents using equilibrium strategies to negotiate with people were generally unsuccessful [17,28]. This prompted us to identify the reasons behind the success of the CS-A agent.

To this end, we compared the behavior of the CS-A agent in HvsCS-A games to that of people in HvsH games in the three countries. We found that in all countries the CS-A agent reached the goal more often than did humans (100% vs 80%).

We therefore conjectured that the high competitiveness of the CS-A played an important role in its success. To support this claim, we found that the benefit for the CS-A agent from the commitment offers that were accepted by people was significantly higher than that reached by people when accepting such offers from other people. In fact, in the commitment offers made by the CS-A, almost all the points that the SP player received from the CS-A reaching its goal (145 points out of 150) went to the CS-A agent.[7]

The competitiveness of the CS-A agent's offers are similar in spirit to (sub-game perfect) equilibrium offers made in the classic two-player ultimatum game. However, people do not generally accept such offers because they are perceived to be unfair. In contrast, in our study, 68.7% of the HvCS-A games ended with people accepting offers that are significantly more beneficial to the agent than to themselves. Interestingly, the average number of rounds in the HvsCS-A games was significantly longer than the HvsH and HvsSPa games (10.12 rounds in HvsCS vs 5.11 rounds in HvsH vs 3.9 in HvsSPa). A possible explanation is that people were aware of the disadvantage to them when accepting the CS-A agent's offers and therefore it took a greater number of rounds for them to accept the CS-A's unfair offers. Nevertheless, most of them eventually accepted the unfair offers with almost no benefit.

A possible explanation for the fact that people were willing to accept unfair offers is the inherent competition between the SPs in the game. This is supported by experiments reported in Roth et al. [29], who compared related two-person ultimatum game and multi-player market environments in Israel, Japan, the United States and Yugoslavia. In both environments, subjects played several times, each time with different people. Both the market and ultimatum game environments chosen for their experiment have an extreme sub-game perfect-equilibrium, in which one player receives all of the benefits from the transaction. The market outcomes converged to the equilibrium in all the countries; however the ultimatum outcomes were different from the equilibrium predictions in all the countries. These differences cannot solely be attributed to the repeated interaction, as in both settings, the subjects played the relevant game repeatedly. Rather, one of the explanations is that the competitive pressure toward equilibrium in the market overwhelms any fairness factors in the players' preferences. Thus, we may hypothesize that equilibrium agents may be beneficial in the presence of competition. Despite the success of equilibrium agents playing the CS role, our findings that the SP-A agent was outperformed by people, and other experiments where agents and humans interact in competitive environments such as auctions [7] reveal that the specific context of the settings plays an important role in equilibrium agents' success. Thus categorizing a priori the type of strategies that will succeed in complex environment is still a challenge.

We showed that it was possible to improve the performance of the SP agent by using a data-driven heuristic for finding preferred offers that are likely to be accepted by the CS player while providing a higher benefit to the SP agent than the preferred offers made by the SP-RAP agent. It may also be possible to improve the CS-A agent. In particular, we observed a few games where the human CS player gained more than 500 points while the average score of the CS-A agent was 387.35. The strategy of the human players that obtained a higher score was to first propose non-commitment offers to one of the SPs and get as many chips from him as possible. Only then did the human who played the CS role made a commitment offer to the other side. The risk in such a strategy is that the other SP may not accept the commitment offer after observing the behavior of the CS agent toward the first SP. More experiments should be run to collect data on the SPs' behavior in such situations.

Another possible improvement of the CS-A's score is to increase the acceptance rate of its commitment offers. As mentioned above, only 63% of the HvsCS-A games ended with an SP human player accepting the commitment offer. A possible way to increase the acceptance rate is to decrease the competitiveness of the offers over time, making the later offers more beneficial for the SP human player. Clearly, decreasing the competitiveness of the offers also reduces the score of the CS-A agent. Therefore, again, being able to predict people's willingness to accept such competitive commitment offers over time is essential for the success of such a modified SP agent. We believe that such a predictor will be country dependent since the percentage of HvsCS-A games that ended with a commitment offer in China was only 0.4 while it was 0.66 and 0.8 in Israel and USA, respectively.

The equilibrium definition breaks ties towards the SP_g player when making offers (Definition 2) and breaks ties towards the SP_y player when moving according to the preferred path (Definition 5). We broke ties in this way in order to provide each service provider agent with a fair opportunity for an advantage in the game, rather than consistently favoring one of the service providers at all decision points. From the theoretical point of view we can break ties arbitrarily but keep all the definitions consistent. The correctness of Theorem 1 will not be affected. In particular, in Definition 2 we can break ties toward the SP_y player (instead of the SP_g player), when making offers. In Theorem 1, in the strategy of CS player (item 1) needs to change so that the CS player breaks ties toward the SP_y player instead of the SP_g player when making

[7] A typical offer was 40 chips to the CS-A (33 red chips and 7 purple chips) in return for 11 gray chips to the SP.

offers. Note that this change would also change the negotiation strategies used in equilibrium on the board game used in our study.

Lastly, we deal with the question of how our model adapts to increasing the number of players. We limited the set of participants to two service providers and a single customer, the minimal number of participants to guarantee competition in a market setting, which also allowed to easily run lab studies with people. We can generalize our model to more than two service providers by changing the SP_i preferred offer (Definition 3) to reasoning about all of the other service providers. This is shown in the revised Definition 6 below for SP_i preferred offers:

Definition 6 *(SP_i preferred offer for n service providers).* Let $\{SP_1, \ldots, SP_n\}$ be a set of service providers. The set of SP preferred offers at state s, $\{\hat{O}^-_{s,1}, \ldots, \hat{O}^-_{s,n}\}$ satisfy the following conditions:

1. $Eu_{cs}(s, \hat{O}^-_{s,i}) \geq Eu_{cs}(s, P^*_s)$, $Eu_i(s, \hat{O}^-_{s,i}) > u_i(s)$ and for all $j \neq i$, $Eu_i(s, \hat{O}^-_{s,i}) \geq Eu_i(s, \hat{O}^-_{s,j})$. If there is no offer that satisfies this condition we set $\hat{O}^-_{s,i} = (\emptyset, \emptyset)$.
2. For any offer O involving an SP_i player that satisfies the constraints in condition (1), if $Eu_i(s, O) > Eu_i(s, \hat{O}^-_{s,i})$ then there exists a service provider $SP_j, i \neq j$ such that either $Eu_{cs}(s, O) < Eu_{cs}(s, \hat{O}^-_{s,j})$, or $Eu_{cs}(s, \hat{O}^-_{s,j}) = Eu_{cs}(s, O)$ and ties were broken in favor of SP_j.

The model also extends to the case in which there are multiple customers, and the chip set of each service provider allows it to make a commitment offer simultaneously to each customer during the negotiation phase. We only need to update the SP_i preferred offer of Definition 3 to reason about making an offer to each of the customer players, rather to just one customer player. The alternative case in which it is not possible for an SP_i player to make commitment offers to all customers is more complicated, and we leave it to future work.

7. Conclusions

This paper studied three-player contract games consisting of human and computer players. We defined a new game that comprises three players, two service providers and one customer. The service providers compete to make repeated contract offers to the customer consisting of resource exchanges in the game. The customer can join and leave contracts at will.

The first contribution of the paper was the design of a "contract game" which is analogous to a market setting in which participants need to reach agreements on contracts over time. The second contribution was providing sub-game Perfect equilibrium strategies for each of the players. Specifically, because service providers compete over the customer player, the contracts proposed by both service providers and customers are highly beneficial contracts to the customer, but require a commitment from the customer that would prevent it from signing a contract with the other service provider.

We evaluated computer agents that use the equilibrium strategies in extensive empirical studies in three different countries, the U.S., Israel and China. We ran several configurations in which two human participants played a single agent participant in various role configurations in the game. The third contribution is showing via our extensive experiments that the computer agent using equilibrium strategies for the customer role was able to outperform people playing the same role in all three countries. In particular, the customer agent made significantly more commitment type proposals than people, and requested significantly more chips from service providers than did people. In contrast, the computer player in the role of the service provider was not able to outperform people due to the fact that people may not behave rationally in the game. However, our forth contribution is that introducing a simple model that captures the bounded rationality of people playing the customer role into the formal model of the contract game led to an automated agent that plays well the service provider role.

These results indicate that in some settings it is possible to build agents that interact proficiently with people using equilibrium strategies. In future work, we mean to use machine learning models to predict how people behave in the game and to use this insight to improve the performance of the computer agents. We will need to collect significantly more data instances of people's play for this purpose. In addition, we mean to conduct more experiments with different board games that vary the dependency relationships between the players, and increase the number of service providers and customers.

Acknowledgements

This work was supported in part by the Israel Science Foundation (grant No. 1488/14) and EU FP7 FET SmartSociety project under Grant agreement No. 600854.

Appendix A

In this section we show how the preferred commitment offer for the SP-A agent is realized in the board game of Fig. 1.

Proposition 3 *(Preferred commitment offer for SP-A agent). The preferred commitment offer of the SP-A agent in the board game of Fig. 1 was to offer 33 red chips and 7 purple chips) in return for 11 gray chips.*

The SP_i preferred offers at the initial state s_1, $\hat{O}^-_{1,i}$ are defined iteratively. At iteration 0, the preferred offer depends on the expected utility $Eu_{cs}(1, P^*_1)$ for the CS player given the preferred path P^*_1. The definition of P^* is recursive. Therefore we show how to compute this offer using backward induction.

1. Suppose that the CS player is located one square before the yellow goal, no offer was accepted until this point, and it is the CS player's turn to make an offer. We denote this as state s_l. If no offer will be accepted at this state, then the CS player will go to the yellow goal. This is the base case of the recursion. When reaching the goal, the CS player will have 15 chips in its possession, and its score will be $150 + 15 \cdot 5 = 225$, the SP_y's score will be $40 \cdot 5 + 150 = 350$ and the SP_g's score will be $40 \cdot 5 = 200$.
2. Given the definition of CS preferred offer the CS will give the offer $\hat{O}_{l,g}$ where it gives 4 yellow for 7 pink and 26 reds to SP_g. The SP_g score will be $150 + 11 \cdot 5 = 205$. The CS players' score will be $(24 - 8 - 7 - 4 - 2 + 33) \cdot 5 + 150 = 330$; the SP_y's score be 200. Note that it is a commitment offer (Definition 4). Otherwise, after the acceptance of the offer the CS will go to G_y leaving SP_g with 55 points.
3. In the prior step, denoted $l - 1$, it is the turn of the SPs players to make offers. $SP_{l-1,y}$ should yield at least 330 to the CS player (see condition 1) and at least 200 points to the SP_y (the score for SP_y in the initial state). The best such offer for SP_y getting 11 grays and sending him 32 reds $(24 - 11 - 9 + 32)$ which will yield the SP_y the score of $150 + 19 \cdot 5 = 245$. For SP_g the only such offer to SP_g is the offer $\hat{O}_{l,g}$. The CS player prefers the SP_y over the SP_g player because it maximizes the sum of the scores.
4. Using the same line of reasoning, we can show that preferred offers for all states on the yellow path leading to the SP_y player up to the yellow square near the "small" purple bridge are as follows: When it is the CS player turn to make an offer, the CS preferred offer will be $\hat{O}_{l,g}$ (and the commitment is towards to SP_g). When it is the turn of the SP players to make an offer, the SP preferred offer will be $\hat{O}^-_{l-1,y}$ (and the commitment is towards the SP_y).
5. Consider the state s_{b1} where the CS player is located in the fourth square on the yellow path (near the purple "bridge") and suppose it is the CS player to make an offer and it must move. If its offer is rejected the best path is to continue on the yellow path $Eu_{cs}(s_{b1}, P^*_{b1}) = 330$, $Eu_y(s_{b1}, P^*_{b1}) = 245$ and $Eu_y(s_{b1}, P^*_{b1}) = 200$. The CS preferred offer, $\hat{O}_{b1}$ is sending 7 yellow chips to SP_g and getting in return 3 purple and 33 red chips which will be accepted by the SP_g player. The expected utility of the CS player will be $(24 - 4 - 7 - 3 - 6 + 36) \cdot 5 + 150 = 40 \cdot 5 + 150 = 350$ and SP_g's score will be 205. SP_y's score will be 200.
6. In the state prior to s_{b1} the SPs will make offers. The preferred offer for SP_g will be $\hat{O}_{b1}$ since this is the only offer that will yield the SP_g more than 200 points and at least 350 to the CS player. However, SP_y's offer will be getting 11 gray chips from the CS player and sending back 33 red chips and 3 purple in return. The CS's score will be $(24 - 9 - 11 + 36) \cdot 5 + 150 = 350$ and SP_y's score will be $(40 + 11 - 36) \cdot 5 + 150 = 15 \cdot 5 + 150 = 225$. Breaking ties toward the offer that maximizes the sum of the expected utility leads to choosing $\hat{O}^-_y$.
7. Going back, the prior interesting state, s_{b2} is when the CS player is located in the second square on the yellow path and it is its turn to make an offer and it must move. The offer should yield at least 350 to CS and 205 to SP_g or 245 to SP_y. The best such offer to the CS player, $\hat{O}_{b2}$ is sending 9 yellow chips to SP_g and getting in return 7 purple and 31 red chips. The CS score will be $(24 - 9 - 2 - 3 - 6 + 38) \cdot 5 + 150 = 360$ and the expected utility of SP_g is 205.
8. In a prior state when it is the SPs to make others, and similar to s_{b1} also here SP_g preferred offer is $\hat{O}_{b2}$. The preferred offer of SP_y is getting 11 gray chips and sending 33 red chips and 5 purple chips in return. Breaking ties in favor of the offer that maximize the sum of the expected utility SP_y's offer will be accepted. The expected utility for the CS will be $(24 - 11 - 9 + 38) \cdot 5 + 150 = 360$ and SP_y's expected utility will be 215 and that of SP_g will be 200.
9. Using the same line of reasoning these will be the offers also when the CS is in the first square of the yellow path.
10. Now we are ready to compute the expected utility $Eu_{cs}(2, P^*_2)$ where $s_2 = Res(1, P^*_1)$. By the definition of the preferred path, P^*_1 is the empty set, therefore the CS player must move at s_2. Since we break ties toward SP_y, the CS player will move one square toward the yellow goal. In this case, from item 8 we have that $Eu_{cs}(2, P^*_2) = 360$, $Eu_y(2, P^*_2) = 215$ and $Eu_g(2, P^*_2) = 200$. We need now to compute $\hat{O}_2$. According to the definition of the CS preferred offer $\hat{O}_2$ should yield strictly higher score to the SP as the expected utility from P^*_2 and at least as the expected score from P^*_2 to the CS player. The best offer CS player in this case is made toward SP_g where SP_g is $\hat{O}_2$ which sends 33 red chips and 7 purple chips for 11 yellow chips. This yields 370 points to the CS player, 205 to SP_g.
11. Now we are ready to compute $\hat{O}^-_1$. $\hat{O}^-_{1,g}$ is equal to $\hat{O}_2$. And similarly, for SP_y, the offer is $\hat{O}^-_{1,y}$ is sends 33 red chips and 7 purple chips for 11 gray chips. Breaking ties in favor of SP_g will lead to the acceptance of its agreement.

References

[1] B. An, Automated Negotiation for Complex Multi-Agent Resource Allocation, PhD Thesis, University of Massachusetts, Amherst, 2001.
[2] B. An, N. Gatti, V. Lesser, Extending alternating-offers bargaining in one-to-many and many-to-many settings, in: Proceedings of the 9th IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Sep. 2009, pp. 423–426.
[3] B. An, V. Lesser, K.M. Sim, Decommitment in multi-resource negotiation, in: Proceedings of the International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2008, pp. 1553–1556.
[4] A. Azaria, Y. Auman, S. Kraus, Automated agents for reward determination for human work in crowdsourcing applications, Auton. Agents Multi-Agent Syst. 28 (6) (2014) 934–955.

[5] Y. Bakos, The emerging role of electronic marketplaces on the internet, Commun. ACM 41 (8) (Aug. 1998) 35–42.
[6] J. Dang, M. Huhns, An extended protocol for multiple-issue concurrent negotiation, in: Proceedings of the 20th National Conference on Artificial Intelligence, 2005, pp. 65–70.
[7] R. Das, J. Hanson, J. Kephart, G. Tesauro, Agent–human interactions in the continuous double auction, in: Proceedings of the International Joint Conference on Artificial Intelligence, 2001, pp. 1169–1178.
[8] S. Ficici, A. Pfeffer, Modeling how humans reason about others with partial information, in: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems, Volume 1, 2008, pp. 315–322.
[9] S.G. Ficici, A. Pfeffer, Simultaneously modeling humans' preferences and their beliefs about others' preferences, in: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems, Volume 1, 2008, pp. 323–330.
[10] Y. Gal, B. Grosz, S. Kraus, A. Pfeffer, S. Shieber, Agent decision-making in open mixed networks, Artif. Intell. 174 (18) (2010) 1460–1480.
[11] Y. Gal, S. Kraus, M.J. Gelfand, H. Khashan, E. Salmon, An adaptive agent for negotiating with people in different cultures, ACM Trans. Intell. Syst. Technol. 3 (1) (2012), Article No. 8.
[12] Y. Gal, A. Pfeffer, Modeling reciprocity in human bilateral negotiation, in: Proceedings of the National Conference on Artificial Intelligence, AAAI, 2007.
[13] N. Gatti, F.D. Giunta, S. Marino, Alternating-offers bargaining with one-sided uncertain deadlines: an efficient algorithm, Artif. Intell. 172 (8–9) (2008) 1119–1157.
[14] G. Haim, Y. Gal, M. Gelfand, S. Kraus, A cultural sensitive agent for human–computer negotiation, in: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, 2012, pp. 451–458.
[15] K. Hindriks, D. Tykhonov, Opponent modelling in automated multi-issue negotiation using Bayesian learning, in: Proceedings 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems, AAMAS, 2008.
[16] A. Kalandrakis, A three-player dynamic majoritarian bargaining game, J. Econ. Theory 116 (2) (2004) 294–322.
[17] S. Kraus, P. Hoz-Weiss, J. Wilkenfeld, D. Andersen, A. Pate, Resolving crises through automated bilateral negotiations, Artif. Intell. 172 (1) (2008) 1–18.
[18] S. Kraus, D. Lehmann, Designing and building a negotiating automated agent, Comput. Intell. 11 (1) (1995) 132–171.
[19] C. Li, J.A. Giampapa, K. Sycara, Bilateral negotiation decisions with uncertain dynamic outside options, in: Special Issue on Game-Theoretic Analysis and Stochastic Simulation of Negotiation Agents, IEEE Trans. Syst. Man Cybern., Part C 36 (1) (Jan. 2006) 31–44.
[20] R. Lin, S. Kraus, J. Wilkenfeld, J. Barry, Negotiating with bounded rational agents in environments with incomplete information using an automated agent, Artif. Intell. 172 (6–7) (2008) 823–851.
[21] S. Mascarenhas, N. Degens, A. Paiva, R. Prada, G.J. Hofstede, A. Beulens, R. Aylett, Modeling culture in intelligent virtual agents, Auton. Agents Multi-Agent Syst. (2015) 1–32.
[22] S. Mascarenhas, R. Prada, A. Paiva, G.J. Hofstede, Social importance dynamics: a model for culturally-adaptive agents, in: Intelligent Virtual Agents, Springer, 2013, pp. 325–338.
[23] G. Means, D.M. Schneider, Metacapitalism: The e-Business Revolution and the Design of 21st-Century Companies and Markets, Foreword by J.J. Schiro, John Wiley & Sons, Inc., 2000.
[24] J. Mell, G.M. Lucas, J. Gratch, A. Rosenfeld, Saying yes! The cross-cultural complexities of favors and trust in human–agent negotiation, in: 2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015, Xi'an, China, September 21–24, 2015, 2015, pp. 194–200.
[25] T.D. Nguyen, N.R. Jennings, A heuristic model of concurrent bi-lateral negotiations in incomplete information settings, in: Proceedings of the 18th International Joint Conference on Artificial Intelligence, 2003, pp. 1467–1469.
[26] E. Nouri, D. Traum, Cross cultural report of values and decisions in the multi round ultimatum game and the centipede game, in: 4th International Conference on Cross-Cultural Decision Making, AHFE, 2015.
[27] E. Nouri, D.R. Traum, Cultural differences in playing repeated ultimatum game online with virtual humans, in: Proceedings of the 47th Hawaii International Conference on System Sciences, 2014, pp. 1213–1220.
[28] N. Peled, Y.K. Gal, S. Kraus, A study of computational and human strategies in revelation games, Auton. Agents Multi-Agent Syst. 29 (1) (2015) 73–97.
[29] A.E.R.V. Prasnikar, M. Okuno-Fujiwara, S. Zamir, Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study, Am. Econ. Rev. (1991) 1068–1095.
[30] A. Rosenfeld, I. Zuckerman, E. Segal-Halevi, O. Drein, S. Kraus, NegoChat-A: a chat-based negotiation agent with bounded rationality, Auton. Agents Multi-Agent Syst. 30 (1) (2016) 60–81.
[31] A. Rubinstein, Perfect equilibrium in a bargaining model, Econometrica 50 (1) (1982) 97–109.
[32] T. Sandholm, N. Vulkan, Bargaining with deadlines, in: Proceedings of the 16th National Conference on Artificial Intelligence, AAAI, 1999, pp. 44–51.
[33] T. Sandholm, Y. Zhou, Surplus equivalence of leveled commitment contracts, Artif. Intell. 142 (2002) 239–264.
[34] A. van Wissen, Y. Gal, B. Kamphorst, V. Dignum, Human–agent team formation in dynamic environments, Comput. Hum. Behav. 28 (2012) 23–33.
[35] X.S. Zhang, M. Klein, I. Marsá-Maestre, Complex contract negotiation with structured search and agenda management, in: Proceedings of the 29th National Conference on Artificial Intelligence, 2005, pp. 1507–1514.