Automated Strategies for Determining Rewards for Human Work

Amos Azaria[1] and **Yonatan Aumann**[1] and **Sarit Kraus**[1,2]

[1]Department of Computer Science, Bar-Ilan University, Ramat Gan 52900, Israel
[2]Institute for Advanced Computer Studies, University of Maryland, MD 20742
{azariaa1,aumann,sarit}@cs.biu.ac.il

Abstract

We consider the problem of designing automated strategies for interactions with human subjects, where the humans must be rewarded for performing certain tasks of interest. We focus on settings where there is a single task that must be performed many times by different humans (e.g. answering a questionnaire), and the humans require a fee for performing the task. In such settings, our objective is to minimize the average cost for effectuating the completion of the task. We present two automated strategies for designing efficient agents for the problem, based on two different models of human behavior. The first, the Reservation Price Based Agent (RPBA), is based on the concept of a reservation price, and the second, the No Bargaining Agent (NBA), uses principles from behavioral science. The performance of the agents has been tested in extensive experiments with real human subjects, where NBA outperforms both RPBA and strategies developed by human experts.

Introduction

The problem of motivating people to complete tasks is important in many different situations, from encouraging children to do their homework, motivating people to exercise or making sure a worker delivers his product on time. Providing monetary rewards upon the completion of a task is one of the most common mechanisms used to ensure a person's satisfactory performance. The challenge in using a monetary mechanism is the tradeoff between providing high rewards that increase the probability of task completion and the desire to minimize costs. In this paper we consider automated strategies for determining an efficient reward structure for facilitating the completion of a certain task by human subjects. We concentrate on settings where there is a single task that has to be performed many times by different humans. Such settings include: answering a questionnaire - where it is necessary that the same questionnaire be answered by many subjects; completing a course or a program; participation in an experiment; and more.

In such settings, our goal is to design automated agents that successfully interact with the humans, thereby facilitating the completion of the task by the human subjects. In order to entice subjects to complete the task, the agent can

offer monetary rewards. In addition, there may be some upfront cost to recruit subjects. Given this cost model, our goal is to develop agents that minimize the cost for effectuating the completion of some predefined number of tasks.

In addition to the basic setting, we consider an extension to the model where the task at hand is composed of a series of *milestones*. Conceptually, the milestones may correspond to separate sections of a long questionnaire, classes of a course or stages of an experiment. In such cases it may be necessary to provide a reward for the completion of each milestone separately, but the goal of the agent is still to effectuate the completion of the entire task. Thus, money may be wasted on uncompleted tasks, and the agent needs to anticipate the probability that the human indeed completes the entire task.

The Formal Setting. We consider the following setting. A software agent, which we call *the requester*, has a *task*, T, that can be performed by human subjects. The requester needs the task to be performed some fixed number of times. There is an unlimited stream of human *workers*, each of whom can perform the task, but may require a monetary reward for doing so. To determine the reward, the requester makes an offer to the worker, who can either accept or reject the offer. If accepted, the human worker performs the task, gets the reward, and the requester moves on to the next worker in line. If the offer is rejected, the requester can either make another offer to the same worker or abandon this worker and start interacting with the next worker. However, the worker doesn't know whether a given offer is the last offer or if it will be raised upon rejection. A cost, C_o, is associated with making each offer, and there is a cost of C_c for bringing in a new worker. The process ends when the task is performed the necessary number of times. The objective of the requester is to complete the process while minimizing the total cost. We call this the *Task Completion Game (TCG)*, and seek efficient automated strategies for the requester in such games.

In the *Milestones Task Completion Game (M-TCG)*, the task T is composed of a sequence of k *milestones* $T = \{m_1, m_2, ..., m_k\}$. In order to successfully complete a task, all milestones of the task must be completed *by the same worker*. The worker is rewarded for performing each milestone separately and has the right to leave at any time, even

after getting a reward for some but not all of the milestones. For this version of the game, we consider two possible processes for determining the rewards. In the *stepwise process* the reward for each milestone is determined separately before performing the specific milestone. In the *upfront process* the individual rewards for all milestones are determined in advance (before starting the first milestone). As mentioned, in both cases, the worker may leave at any time, keeping whatever rewards he has accumulated so far.

Thus, in total we consider three settings: the basic TCG (with an indivisible task), and two versions of the M-TCG - stepwise reward determination and upfront reward determination.

Methods and Results. The core premise of this work is that efficient interaction with humans requires a proper understanding and modeling of their behavior. For example, while an equilibrium strategy is theoretically considered the most rational one, agents using such strategies often perform poorly in practice (Peled, Gal, & Kraus 2011; Hoz-Weiss *et al.* 2008). This is because humans commonly do not use equilibrium strategies themselves, hence replying with such a strategy can be sub-optimal. Thus, it is important to develop a good model of the true human behavior, possibly also including psychological factors, in order to optimize the performance of agents interacting with these humans (Gal & Pfeffer 2007; Hindriks & Tykhonov 2008; Oshrat, Lin, & Kraus 2009; Rosenfeld & Kraus 2011; Peled, Gal, & Kraus 2011).

In this paper we develop two agents for the Task Completion Game and the variants thereof: RPBA (*Reservation Price Based Agent*) and NBA (*No Bargaining Agent*), each based on a different model for human behavior. We performed extensive experiments with an actual TCG carried out on Amazon's Mechanical Turk service with 903 people. We measured the performance of our agents in the three different settings. We also compared the performance of these automatically generated agents to the performance of policies manually determined by human experts. The results are described in detail in the following, but the overall outcome is that NBA outperforms both RPBA and the human experts. The performances of all three agents are far superior to that which would result from using the equilibrium strategy. Thus, proper modeling of human behavior is essential in such settings and can significantly increase the performance.

Related Work

Negotiation has been studied extensively (Rubinstein 1982; Raiffa 1982). Most research has considered the archetypal form of negotiation, wherein the sides alternate in making offers (see (Stahl 1972; Di Giunta & Gatti 2006; Lin *et al.* 2008)). This model was extended by (Ramchurn *et al.* 2007) who specify commitments that agents make to each other when engaging in persuasive negotiations using rewards. In practice, however, this is often not the actual dynamic. In many real world situations it is only one side that makes the offers, and the other side merely accepts or rejects them. This is most commonly the case, for example, in the job market. Employers make offers to job candidates, and may improve the offers if rejected by the candidates. Job candidates, on the other hand, most commonly simply accept or reject the offers and are much less likely to make counter offers.

The *Secretary Problem* (see (Ferguson 1989)) is a stopping point problem, where in its simplest form there is a secretary opening with n candidates. The candidates are interviewed in random order and a rejected candidate cannot be recalled. The interviewer is interested in maximizing his probability for choosing the best candidate. It has been shown that the best policy is to reject the first $1/e$ applicants and then choose any applicant better than all applicants seen so far. Using this policy, the interviewer will find the best candidate with a probability of $1/e$. Our problem differs from the secretary problem by allowing bargaining with each candidate. We are not interested in the best (or cheapest) candidate rather minimizing the expected cost.

Sandholm and Gilpin (2006) study sequences of take-it-or-leave-it offers, where a seller proposes a sequence of offers to a set of buyers and each buyer in turn either accepts the offer, pays that amount, obtains the good and ends the game, or rejects the offer and the next buyer receives his own offer. However, in their study the seller must announce the full sequence in advance, leading to a completely different strategy than we intend to study.

The ultimatum game (see (Gaith, Schmittberger, & Schwarze 1982)) is a well known game with two players in which a proposer suggest a way to divide a fixed sum, and a responder may either accept the proposal and each player will get its share, or reject it in which case both the proposer and the responder receive nothing. Although the sub-game perfect equilibrium (SPE) in the ultimatum game is that the proposer retains nearly the entire amount, it has been shown that people do not follow this SPE (see (Cameron 1999; Katz & Kraus 2006)).

Greezy et al. (2003) study the reverse ultimatum game, which is identical to the ultimatum game with one exception: any time the responder rejects an offer, the proposer may suggest another offer, as long as the new offer is strictly higher than the previous one. This change inverts the SPE, leaving the responder with nearly the entire sum and the proposer with nearly nothing. Once again, Greezy et al. show that people do not follow this SPE, but do not suggest any strategy for the proposer.

Woolley et al. (2010) study "collective intelligence for groups of people. They define a collective intelligence factor and state that this factor appears to depend both on the composition of the group and on factors that emerge from the way group members interact when they are assembled. Although we intend to design a system that interacts with many workers, we concentrate on settings where there is a single task that has to be performed many times by different humans independently, and do not consider collaborative tasks.

Galuscak et al. (2010) provide a survey on criteria which are used to determine the wages of newly hired employees in Europe. In (Walque *et al.* 2009) Walque et al. provide a

formal model of a labor market, including methods for payment and wage negotiation. However, these papers merely study the nature of wage determination and do not provide employers with any strategies for determining appropriate wages.

Agents

We will now describe our agents for the Task Completion game. For completeness, we first determine the equilibrium strategy for this game.

Equilibrium Strategy

In the Task Completion Game, if the worker is aware that the requester must complete the task and knows the maximum offer that can be made by the requester, then the only sub-game perfect equilibrium is where the requester proposes the maximum possible offer and the worker accepts it (and rejects any lesser offer). This result is similar to the sub-game perfect equilibrium in the reverse ultimatum game studied in (Gneezy, Haruvy, & Roth 2003).

It has already been shown by (Gneezy, Haruvy, & Roth 2003) (and confirmed by our own results) that people do not use the SPE strategy in such games. Therefore, an agent performing on behalf of a requester who proposes the maximum offer will obviously perform poorly (unless the maximum offer is unreasonably low). Mal-performance of agents using equilibrium strategies when interacting with people, was shown in (Hoz-Weiss *et al.* 2008; Peled, Gal, & Kraus 2011; Azaria *et al.* 2011). We therefore model human behavior using other approaches.

Reservation Price Agent

Since humans do not use equilibrium strategies, we consider other models for their behavior. The first model we consider is based on the notion of a *Reservation price*. Specifically, the model assumes that for each task (or milestone) each worker has a unique *Reservation price (RP)* that is the amount the worker requires to perform the task; it will accept any offer greater than or equal to this amount and reject any lesser offer. Note that reservation price behavior is not strategic (in the game-theoretic sense) as it does not take into account the considerations of the other side.

We now describe the construction of the reservation price agent for the case of a single task with no milestones. The construction for the setting with milestones is similar in nature but more complex, and its description is left to the full version of this paper.

Reservation Price Elicitation. In order to construct a policy for the agent, we must know the reservation prices of workers, or the distribution thereof. Since we are not given this distribution, we approximate it by sampling a subset of the workers and eliciting the reservation prices by means of a truthful auctioning mechanism.

We consider the following two methods, in which the dominant policy for each player is to bid for his truthful reservation price. The first is the Vickrey auction (Vickrey 1961). A Vickrey auction is a sealed bid auction where each worker submits a bid and the worker with the lowest bid performs the task and is paid the amount requested by the second lowest bid. Allocating tasks using a Vickrey auction was used, for example, in (Sarne & Kraus 2005). The second method is the Becker-DeGroot-Marschak (BDM) mechanism (Becker, DeGroot, & Marschak 1964). In the BDM mechanism, the worker bids for an amount, then a computer selects a number by random, if the random number is lower than the bid, no work is done; however, if the random number is greater than the bid, the worker performs the task and is paid the sum given by the random number.

Noussair et al. (Noussair, Robin, & Ruffieux 2004) compare the Becker-DeGroot-Marschak (BDM) mechanism with the Vickrey Auction and show that although both methods are biased, the Vickrey auction is more effective for eliciting the true reservation price.

We therefore, sample a subset of the workers at random and perform a Vickrey auction with them. In order for workers to have a reasonable chance of winning, we divide the set of workers into subsets, where each worker only competes with the bids offered by the other workers in its subset. For any price x, denote by $p(x)$ the fraction of workers with reservation price x, and $P(x) = \sum_{x' \leq x} p(x')$.

Optimal Policy RPBA Algorithm. Denote the number of different reservation prices by ℓ, let $rp_1 < rp_2 < \cdots < rp_\ell$ be the different reservation prices, and set $rp_0 = -\infty$. Clearly there is no sense in making any offer other than one of the reservation prices. Further note that under the reservation price assumption, for any deterministic policy there is necessarily some reservation price rp_s, such that any worker with a reservation price greater than rp_s will reject all offers it gets, and any worker with a reservation price smaller than or equal to rp_s will accept some offer. Since we do not know s, we test all possible values for s. For each, we compute the optimal policy assuming all and only agents with a reservation price of at most rp_s will accept an offer (as explained below). We then calculate the expected cost for each of these policies and choose the policy with the lowest expected cost.

Suppose we want to accept all and only workers with a reservation price of at most $\leq rp_s$. Clearly, since workers with a reservation price of rp_s need to accept, we must make an offer of at least rp_s. However, since workers with a higher reservation price must reject, we cannot make a higher offer. Thus, the offer rp_i must be made, and it will be the last offer.

For workers with lower reservation prices, it would be best to offer each specific worker its exact reservation price. However, there is a cost for each offer. Thus, there is a tradeoff between not offering much more than necessary and minimizing the number of offers. We determine the optimal offer sequence using dynamic programming, as follows.

We construct a table T of size $s \times s$, as follows. Consider j, i with $1 \leq j \leq i \leq s$. Suppose a worker has rejected the offer rp_{j-1}, entry $T(i, j)$ stores the expected cost of completing the interaction with this worker, under the optimal strategy, with the following constraints:

- The next offer must be rp_i.

- All and only workers with reservation price $\leq rp_s$ accept an offer.

We now show how to fill-in the entries of T. Recall that C_c denotes the cost for calling in a worker and that C_o denotes the cost for proposing an offer. For a given $i \geq j$, let $p_a(i,j)$ be the fraction of workers with a reservation price of at least rp_j who would *accept* offer rp_i:

$$p_a(i,j) = \frac{\sum_{x=rp_j}^{rp_i} p(x)}{\sum_{x \geq rp_j} p(x)}$$

Since the offer rp_s must be made to workers with a reservation price of rp_s, we have

$$T(s,s) = C_o + p_a(s,s) \cdot rp_s \qquad (1)$$

The expected cost of completing the interaction with current worker after *rejection* of the offer rp_i is:

$$c_r(i) = \min_{i' \geq i+1}\{T(i', i+1)\} \qquad (2)$$

For other i, j, we have:

$$T(i,j) = C_o + p_a(i,j) \cdot rp_i + (1 - p_a(i,j))c_r(i) \qquad (3)$$

Using (1), (2) and (3), table T is filled out in descending order of j and i.

The overall expected cost per worker, assuming that all and only workers with reservation prices of, at most, rp_s accept, is:

$$\min_i\{T(i,1)\} + C_c \qquad (4)$$

The expected number of workers with a reservation price of, at most, rp_s is $P(rp_s)$. Thus, if only workers with a reservation price of, at most, rp_s accept, the expected cost until the task is performed is:

$$cost(s) = \frac{\min_i\{T(i,1)\} + C_c}{P(rp_s)}$$

We iterate through all possible values of s to find the optimal one. By saving the indexes found while computing the minimum in (1) and (4), the algorithm outputs a policy for the RPBA (Reservation Price Based Agent), detailing exactly what offers to make at any given stage, including the decision of if and when to abandon the existing worker and seek another one.

The algorithm runs in $O(l^3)$ time and $O(l^2)$ space.

No Bargaining Agent

Experiments with the Reservation Price based agent have produced mixed results. More precisely, it seems that people frequently do not have a fixed price point at which they are willing to perform the task. Rather, the interaction with the requester largely determines the desired price (so long as it is within a reasonable range). In particular, just knowing that the price is negotiable seems to push up the price. We call this *the bargaining effect*. More Formally, suppose a worker's minimum offer that he will accept as a first offer is y. The bargaining effect claims, that the worker is likely to reject this offer (y) if it follows a lower offer already rejected

by the same worker. Indeed, Riley and Zeckhauser (Riley & Zeckhauser 1983) have shown that a seller receives the highest utility when he does not allow any bargaining at all. Our experiments support this observation (see Experimental Evaluation Section). Thus, we develop the *No Bargaining Agent (NBA)*. This agent makes only one offer for each task or milestone (to any given worker) and never makes a second offer if rejected. Again, due to space constraints, we describe the algorithm for the basic TCG and leave the M-TCG settings to the full version.

The best (one and only) offer to make depends on what offers would and would not be accepted, and the distribution thereof. We thus wish to estimate this distribution. We assume that the portion of workers that would accept a first-time offer of x follows a sigmoidal distribution (in x) (Gal & Pfeffer 2007). We approximated the distribution by choosing several points, then sampled a subset of workers and obtained their acceptance fraction for these points (different workers for different points) and interpolated the sigmoid from these values. Let $u(x)$ be the fraction of workers estimated to accept a first-time offer x (out of the entire set of workers).

Finding an optimal policy. Given the acceptance distribution, we wish to find the optimal policy. Due to the no-bargaining policy, once an offer is rejected, NBA does not attempt to make another offer, but calls for the next worker. Therefore, for the no-milestone case, NBA's policy consists of a single offer. With a policy of offering x, the expected cost per worker is:

$$C_c + C_o + x \cdot u(x)$$

The expected number of workers sampled until a worker accepts the offer is $1/u(x)$. Thus, the expected cost per completed tasks is:

$$cost(x) = \frac{C_c + C_o + x \cdot u(x)}{u(x)} \qquad (5)$$

Using this cost as our fitness function, we perform a search to find the optimal offer x.

Experimental Evaluation

All of our experiments were performed using Amazon's Mechanical Turk service (AMT) (Amazon 2010)[1]. Participation in all experiments consisted of a total of 1152 subjects from the USA, of which 53.3% were females and 46.7% were males. The subjects' ages ranged from 18 to 79, with a mean of 33 and median of 27. All subjects were paid 11 cents for participating in the study. Any extra credit gained in the task was given to the subjects as a bonus.

All tasks and milestones were composed of a set of simple puzzles where the subjects were required to find a distinctive shape among other shapes (see Figure 1 for example). The different milestones varied in the number of puzzles needed to be solved by the worker and the number of shapes.

_____

[1]For a comparison between AMT and other recruitment methods see (Paolacci, Chandler, & Ipeirotis 2010).

Figure 1: Example for a single puzzle
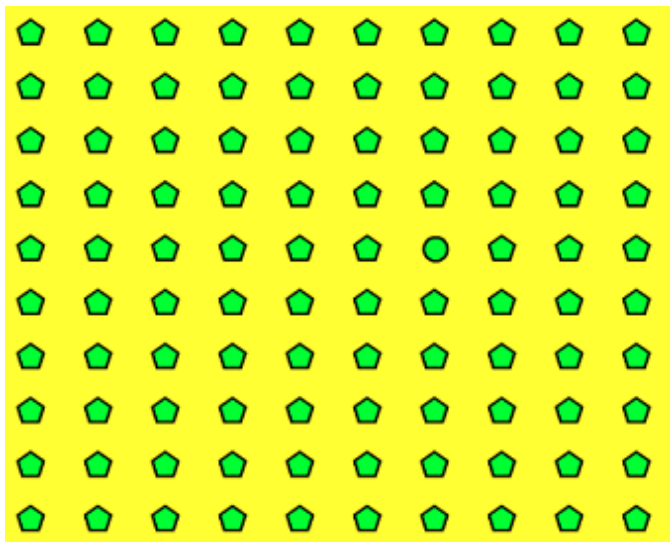
In all of the experiments, all of the subjects were first required to complete a small set of easy puzzles (without being paid any additional amount) to make sure that they understood the nature of the task.

We set C_c (cost for calling a new worker) to 20 cents, and C_o (cost per offer) to 4 cents. In all of the games the subjects were unaware of the full protocol.

Experimental Introductory Overview

The subjects (1152 in total) participated in the following experiments:

- Bargaining Effect Experiment: this experiment tested the basic assumption of the NBA agent, and showed that proposing an offer which is too low may raise the final price. 60 subjects participated in this experiment.

- RPBA Price Elicitation: the RPBA agent uses a Vickrey Auction to elicit the types. 183 subjects participated in the Vickrey Auction.

- NBA Price Elicitation: 225 subject participated in the acceptance distribution learning stage.

- Comparisons Between all Three Agents (RPBA, NBA and EXPERTS): 684 subjects participated in the comparison between the agents among all three different settings. (105 subjects in the basic TCG, 318 in M-TCG stepwise reward determination, and 261 in the M-TCG upfront reward determination.)

The basic TCG was composed of a single milestone (by definition). Both M-TCG with stepwise reward determination and M-TCG with upfront reward determination were composed of five milestones ($k = 5$).

Bargaining Effect Experiment

We first tested the bargaining effect assumption on the Task Completion Game, using the following experiment. We recruited 60 subjects and split them into two groups - the "No Bargaining Group" and the "Bargaining Group" (30 workers in each). Workers in the "Bargaining Group" were first offered 1 cent for performing the task, and, if rejected, were offered 4 cents for the same task. Workers in the "No Bargaining Group" were offered 4 cents for performing the task with no other offer. As in the general case, neither of the two groups knew if they would receive any sequential offers upon rejection.

Table 1 lists the results of the experiment. In the "No Bargaining Group" 24 subjects accepted the 4 cent offer. In the "Bargaining Group" 13 subjects accepted the first offer (of 1 cent), however, among those who rejected the first offer, only a single subject accepted the second offer (of 4 cents). Thus, a total of only 14 subjects accepted the 4 cent offer in this group substantially less than in the No Bargain Group. These results differ significantly using Fisher's exact test ($p < 0.01$).

The psychological basis for such behavior is an interesting question, but the core phenomena was already noted by (Overstreet 1925) and later by (Carnegie 1964) saying (p. 121): "A 'No' response is a most difficult handicap to overcome. When you have said 'No' all your pride of personality demands that you remain consistent with yourself."

Table 1: Bargaining Effect Experiment

group	participants	subjects accepting offer		
		1 cent	4 cents	total
No Bargaining	30	–	24	24
Bargaining	30	13	1	14

In (Ariely, Loewenstein, & Prelec 2006) Ariely et al. demonstrate an effect called anchoring. In their study, subjects were asked if they would be willing to pay as much in dollars as the last two digits in their Social Security Number (SSN) for an item. They were then asked how much they would be willing to pay for that same item. Ariely et al. show a high correlation between the last two digits of the SSN and the amount the subjects were willing to pay. We would like to note that in the bargaining averse experiment above, the second group, who first received an offer of only 1 cent, were actually anchored to a lower offer than the first, therefore according to the anchoring effect, they should have been willing to accept a lower final offer. However, as shown by the results, it seems that (under the settings we used) the bargaining effect is stronger than the anchoring effect.

Manually Designed Agents

In order to evaluate the quality of NBA and RPBA, we also constructed an additional set of agents. We interviewed four people with significant experience as requesters in Mechanical Turk (whom we refer to as experts). After explaining the problem, each expert was requested to compose a policy. We considered the average cost among the four experts as the cost for the manually designed *Experts' agent*.

While one expert used a method similar to NBA, and proposed only a single offer to each worker (for each mile-

stone), the other experts tended to raise their offer (usually only once) upon rejection.

Results

We ran 25 instances of the experiment for each of the agents, i.e. each agent had to accomplish the goal 25 times, except for the manually designed agents which accomplished 8 goals each, leading to a total of 32 goals for all four experts. Participants consisted of 594 subjects.

Tables 2, 3 and 4 provide the results obtained for the basic TCG, the M-TCG with stepwise reward determination and the M-TCG with upfront reward determination, respectively. Results show the average number of workers called, the average number of offers given and the average cost per task performed. The last column shows the acceptance rate - the fraction of accepted offers. As can be seen, NBA required the lowest cost per task among the three agents in all three settings. Combining all experiments together, Figure 2 shows the average performance of the three agents over all three settings we tested. Clearly, NBA performs the best. For the statistical test, since the Experts' Agent performed better than the RPBA agent on average among the three settings, we use it as our baseline and compare the NBA agent to the Experts' Agent. Using the ANOVA test, NBA's better performance is statistically significant ($p < 0.05$).

Table 2: Average per Goal (completion of a single milestone) in Basic TCG

agent	workers	no. offers	cost	accept. rate
RPBA	1.2	1.84	43.88	54.3%
NBA	1.44	1.44	**42.56**	69.4%
EXPERTS	1.22	2	48.13	68.0%

Table 3: Average per Goal (completion of all 5 milestones) in M-TCG, Stepwise Reward Determination

agent	workers	no. offers	cost	accept. rate
RPBA	6.72	24.56	297.88	64.3%
NBA	4.32	12.88	**165.16**	74.2%
EXPERTS	1.31	6.44	184.3	86.1%

Table 4: Average per Goal (completion of all 5 milestones) in M-TCG, Upfront Reward Determination

agent	workers	no. offers	cost	accept. rate
RPBA	4.76	17.00	191.64	64.4%
NBA	3.56	11.32	**140.00**	77.7%
EXPERTS	1.66	6.5	200.97	87.0%

Discussion

The core principle of the NBA agent, which performed best, is the principle of "no bargaining". Although the bargaining effect has been mentioned in the past (both by social scientists and theoreticians) this paper is not only the first to test
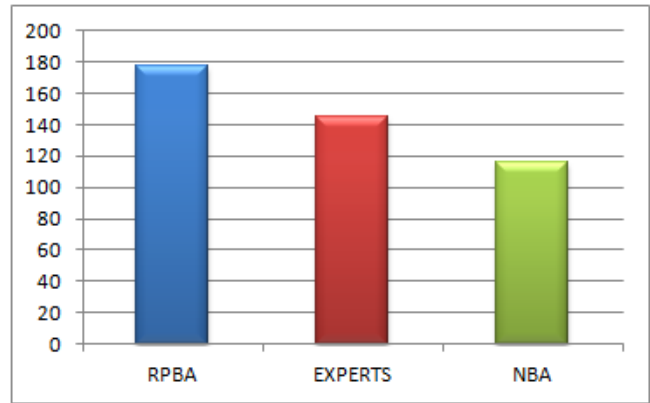

Figure 2: Average cost per goal over all three settings

this effect and its power in practice, but also to assimilate it into an automated agent.

Examining the detailed dynamics of the games exhibits how the willingness of the other agents to bargain harmed their performance. In the basic TCG experiments, there were 11 cases in which a worker rejected some offer and the agent (not NBA) replied with an improved offer. Of these workers, only a single worker ended up completing the task. For the remaining 10, all subsequent offers were rejected, and the cost thereof wasted. In the multiple milestone experiments the effect was even more pronounced. In all, in the M-TCG experiments there were 110 workers that rejected an offer and were offered a better one (not by NBA). Of these 25 accepted one of these future offers. However, only 4 of these 25 ended up completing the entire task. Thus, only 4 of the 110 completed the task, but a substantial cost was invested on the other ones. Thus, bargaining seems to be fruitless for the most part, mostly adding to the expenses and seldom bringing about the desired outcome.

We further show that the NBA's price elicitation method is much more accurate than the RPBA's. In the basic TCG, when predicting the worker's response, the RPBA agent had a Mean Squared Error (MSE) of 0.46, which is extremely high, while the NBA agent had an error of only 0.18. This result suggests that price elicitation by sampling a subset of workers and interpolating a sigmoid, is much more accurate than using a Vickrey Auction for the same purpose. Furthermore, due to the NBA's price elicitation method, it can use any data gained while interacting with humans to refine its believes on human reaction and regenerate a more accurate strategy every interaction.

Also worth noting is that while NBA provided the least overall average cost, the Experts' Agent exhibited a higher acceptance rate and a smaller average number of workers per task completion (see Tables 2-4). This might suggest that in the Task Completion Game people tend to make offers that are too high. The reason for this is a subject for further research, but perhaps it can be linked to the sunk cost effect (see (Arkes & Blumer 1985)), where in order to not appear wasteful in the present, people tend to invest more money in a situation where greater money has already been invested,

regardless of future expected expenses and utility. Also, the starting offer of the Experts' Agents was higher than that of the automated one, perhaps also linked to the human preference not to "fail".

We illustrate the above insights using the simple case of the basic TCG (with a single milestone). Table 5 shows the policies constructed by each of the agents along with their average cost per task completion in the basic TCG. As shown in Table 5, the experts either proposed too many offers (e.g. Expert #4) or proposed an offer which was too high to start with (e.g. Expert #2), leading to costs which were higher than both NBA and RPBA.

Table 5: Reward Determination Policy for Basic TCG

agent	1st offer	2nd offer	3rd offer	4th offer	5th offer	avg. cost
RPBA	15	25	40	50	-	43.9
NBA	8	-	-	-	-	42.6
Expert #1	17	22	-	-	-	48.0
Expert #2	22	-	-	-	-	49.0
Expert #3	10	15	-	-	-	44.5
Expert #4	22	33	36	37	50	51.0

Finally, it is interesting to note the differences in the results of the two variants of milestones-TCG - the upfront reward determination schedule and the stepwise one. Both automated agents (RPBA and NBA) performed better in the upfront schedule than in the stepwise one, with RPBA's cost being $\sim 30\%$ less in the upfront schedule than in the stepwise one. This is the behavior we expected since the commitment on the side of the requester would tend to lower the cost required by the workers. Interestingly, the Experts' Agents performed slightly worse in the upfront schedule. We do not know the reason for this, and further experimentation is necessary in order to validate that such behavior is consistent and to examine the possible reasons.

Future Work

In the future, we suggest considering the following extensions as avenues for possible improvement of NBA:

Expertise and Boredom: When the milestones are relativity similar to each other, a worker may experience either expertise or boredom when performing more than one milestone. If this happens, this may change the necessary reward structure (e.g. a bored worker may need to be paid more). How exactly to incorporate this into NBA is a subject for future research.

Questionnaire Based Learning: In this work we ignored the cost of collecting the preliminary cost data. Both the Vickrey bidding (used in RPBA) and the sigmoid learning (used in NBA) were implemented using real monetary rewards for workers performing the actual tasks in question. The cost of doing so could be substantial, and is justified in cases where the task needs to be subsequently performed many times. However, such upfront cost could be prohibitive in cases where the number of times the task needs to be performed is not that large. In such cases, or

in order to save on learning expenses in general, the costs can be reduced using a questionnaire based approach, whereby workers are only asked what they would require if they had to perform the task, but not really having them do so in practice. Such a method would probably produce less accurate estimates, but would hopefully still provide a good enough approximation. For an additional reduction in costs, subjects chosen to participate in the learning phase may be asked a series of questions regarding each milestone, requiring fewer participants in this phase.

Conclusions

Motivating people to perform tasks is a difficult problem. When a requester needs to recruit a large number of people to perform some task, even human experts find it difficult to design the most effective strategy to minimize costs. Ensuring that people complete complex tasks with several milestones is even more challenging, but necessary in many settings.

In this paper we presented a reward strategy that succeeded in facing these challenges by building a general model of the human response to offers and their attitude toward negotiation. Based on this model, we designed an automated agent that interacts with the humans and is successful in minimizing the costs. The accuracy of the human model has benefited from applying principles adopted from behavioral science. We strongly believe that this methodology is useful, and essential, in designing self-interested agents that interact successfully with people in other domains as well.

Acknowledgments

References

Amazon. 2010. Mechanical Turk services. http://www.mturk.com/.

Ariely, D.; Loewenstein, G.; and Prelec, D. 2006. Coherent arbitrariness: Stable demand curves without stable preferences. *Journal of Business Research* 59(10-11):1053–1062.

Arkes, H. R., and Blumer, C. 1985. The psychology of sunk cost. *Organizational Behavior and Human Decision Processes* 35(1):124 – 140.

Azaria, A.; Rabinovich, Z.; Kraus, S.; and Goldman, C. V. 2011. Strategic information disclosure to people with multiple alternatives. In *Proc. of AAAI*.

Becker, G. M.; DeGroot, M. H.; and Marschak, J. 1964. Measuring utility by a single-response sequential method. *Behav Sci* 9(3):226–232.

Cameron, L. A. 1999. Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry* 37(1):47–59.

Carnegie, D. 1964. *How To Win Friends And Influence People*. Simon and Schuster.

Di Giunta, F., and Gatti, N. 2006. Bargaining over multiple issues in finite horizon alternating-offers protocol. *Annals of Mathematics and Artificial Intelligence* 47:251–271.

Ferguson, T. S. 1989. Who Solved the Secretary Problem? *Statistical Science* 4(3):282–289.

Gaith, W.; Schmittberger, R.; and Schwarze, B. 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3:367–388.

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

Galuscak, K.; Keeney, M.; Nicolitsas, D.; Smets, F. R.; Strzelecki, P.; and Vodopivec, M. 2010. The determination of wages of newly hired employees - survey evidence on internal versus external factors. Working Paper Series 1153, European Central Bank.

Gneezy, U.; Haruvy, E.; and Roth, A. 2003. Bargaining under a deadline: evidence from the reverse ultimatum game. *Games and Economic Behavior* 45:347–368.

Hindriks, K., and Tykhonov, D. 2008. Opponent modelling in automated multi-issue negotiation using bayesian learning. In *Proceedings 7th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 331–338.

Hoz-Weiss, P.; Kraus, S.; Wilkenfeld, J.; Andersend, D. R.; and Pate, A. 2008. Resolving crises through automated bilateral negotiations. *Artificial Intelligence journal* 172(1):1–18.

Katz, R., and Kraus, S. 2006. Efficient agents for cliff-edge environments with a large set of decision options. In *AAMAS*, 697–704.

Lin, R.; Kraus, S.; Wilkenfeld, J.; and Barry, J. 2008. Negotiating with bounded rational agents in environments with incomplete information using an automated agent. *Artif. Intell.* 172(6-7):823–851.

Noussair, C.; Robin, S.; and Ruffieux, B. 2004. Mrevealing consumers' willingness-to-pay: A comparison of the bdm mechanism and the vickrey auction. *Journal of economic psychology* 25:725–741.

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

Overstreet, H. 1925. *Influencing human behavior*. The People's institute publishing.

Paolacci, G.; Chandler, J.; and Ipeirotis, P. G. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5(5).

Peled, N.; Gal, Y.; and Kraus, S. 2011. A study of computational and human strategies in revelation games. In *Ninth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*.

Raiffa, H. 1982. *The art and science of negotiation / Howard Raiffa*. Belknap Press of Harvard University Press, Cambridge, Mass. :.

Ramchurn, S. D.; Sierra, C.; Godo, L.; and Jennings, N. R. 2007. Negotiating using rewards. *Artif. Intell.* 171:805–837.

Riley, J., and Zeckhauser, R. 1983. When to haggle, when to hold firm. *The Quarterly Journal of Economics* 98(2):267–289.

Rosenfeld, A., and Kraus, S. 2011. Using aspiration adaptation theory to improve learning. In *Ninth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*.

Rubinstein, A. 1982. Perfect equilibrium in a bargaining model. *Econometrica* 50(1):97–109.

Sandholm, T., and Gilpin, A. 2006. Sequences of take-it-or-leave-it offers: Near-optimal auctions without full valuation revelation. *Autonomous Agents and Multiagent Systems (AAMAS)*.

Sarne, D., and Kraus, S. 2005. Solving the auction-based task allocation problem in an open environment. In *AAAI*, 164–169.

Stahl, I. 1972. Bargaining theory. *Stockholm School of Economics*.

Vickrey, W. 1961. Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance* 16(1):8–37.

Walque, G. d.; Pierrard, O.; Sneessens, H.; and Wouters, R. 2009. Sequential bargaining in a new keynesian model with frictional unemployment and staggered wage negotiation. Working Paper Research 157, National Bank of Belgium.

Woolley, A. W.; Chabris, C. F.; Pentland, A.; Hashmi, N.; and Malone, T. W. 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science* 330(6004):686–688.