Optimal Incremental Preference Elicitation during Negotiation

Tim Baarslag and Enrico H. Gerding
Agents, Interaction and Complexity Group
University of Southampton
SO17 1BJ, Southampton, UK
{T.Baarslag, eg}@soton.ac.uk

Abstract

The last two decades have seen a growing interest in the development of automated agents that are able to negotiate on the user's behalf. When representing a user in a negotiation, it is essential for the agent to understand the user's preferences, without exposing them to elicitation fatigue. To this end, we propose a new model in which a negotiating agent may incrementally elicit the user's preference *during* the negotiation. We introduce an optimal elicitation strategy that decides, at every stage of the negotiation, how much additional user information to extract at a certain cost. Finally, we demonstrate the effectiveness of our approach by combining our policy with well-known negotiation strategies and show that it significantly outperforms other elicitation strategies.

1 Introduction

Negotiation is a frequently used and important process by which different parties can reach a mutually acceptable agreement [Lomuscio *et al.*, 2001]. At the same time, it is also a time-consuming and expensive activity that humans often find challenging and stressful [Fatima *et al.*, 2014]. Software agents can help alleviate some of the difficulties involved in negotiation by *representing* the user in an automated manner. For example, the agent may provide negotiation support in highly complex negotiation domains, such as purchasing a supercomputer or conducting union negotiations.

To do so effectively, the agent needs to obtain an accurate user model through preference elicitation, which typically precedes the agent negotiation phase. However, in many realistic settings, extracting the necessary preference information is arduous and costly, as multiple interactions with the system can result in user displeasure and bother [Buffett *et al.*, 2004; Fleming and Cohen, 2004]. In addition, the space of negotiation outcomes is often too large to elicit in its entirety.

Until recently, the challenge of an agent not fully knowing its own preferences and developing strategies to address the tradeoff between negotiation outcome and effort expended in elicitation has received very little attention [Boutilier, 2002]. The work by Chajewska et al. [1998; 2000] provides one of the first starting points to this problem, but their solution method often deals with computationally intractable decision procedures. Costly preference elicitation has also been studied in the setting of auctions; notably by Conen and Sandholm [2001] and Parkes [2005]. These works are primarily aimed at designing *mechanisms* that can avoid unnecessary elicitation. Costly preference elicitation may alternatively be cast as a problem in which agents have to allocate costly computational resources to *compute* their valuation [Larson and Sandholm, 2001], but this work focuses on interactions *between* different strategies.

To address the challenge of formulating an effective negotiation and elicitation strategy, we propose a novel and efficient approach whereby the agent incrementally asks the user for information *during* the negotiation process. This reduces time spent on the preference elicitation phase by minimizing queries to the user, and allows the agent to interactively extract the most valuable information from the user at the most relevant time.

Our main contribution is a *generic*, *optimal*, and *efficient* elicitation strategy for a negotiation agent. The elicitation method is generic in that it can be combined with any existing utility-based bidding strategy. It is optimal in the sense that, during the deliberation cycle, the agent will optimally choose whether or not to elicit further information from the user, taking into account both the elicitation costs (i.e. user bother) and the incremental learning effect of any subsequent information gain. In doing so, the strategy will also take into consideration the likelihood that any offer will be accepted by the opponent. Lastly, we show our algorithm is $O(n \log n)$ efficient, which provides a considerable improvement over the naïve approach. As a second contribution, we show in an experimental setting that our elicitation method outperforms benchmark approaches when coupled with existing, well-known bidding strategies, regardless of the user elicitation costs.

2 Problem Description

We consider a setting where the agent is negotiating with an opponent on behalf of a user who is unwilling or unable to fully specify their preferences. Hence, for each possible offer, the agent has two types of uncertainty: (1) whether the opponent will accept the offer (the *opponent* model), and (2) the user's utility of an offer (the *user* model). We assume that the agent has prior information about both the opponent model and user model,[1] which is updated during negotiation.

The opponent model is updated through the offers that are exchanged with the opponent. At the same time, the agent is allowed to approach the user if it believes it requires additional information to improve the user model.

Given this, at every point in the negotiation, the agent needs to decide whether to ask the user for more information, whether to submit or accept an offer, or whether to end the negotiation. Each of these actions has potential benefits and costs. Specifically, asking the user for information incurs a user bother cost. Hence, the agent should take into account whether this outweighs potential benefits compared to the already known information. The agent's decision problem is especially challenging because we allow *incremental* elicitation: that is, when the extracted preferences of a potential offer prove disappointing, the agent is allowed to continue the elicitation process. Therefore, the agent should balance the value of the current negotiation state with performing one or more elicitation steps. In addition, the agent should take into consideration the likelihood that the opponent is going to accept the offer, as this directly affects the benefits of requesting more precise information.

As well as asking the user, the agent should be careful when accepting an offer, as this forgoes the opportunity to explore further offers. Figure 1 provides an overview of the interaction between the opponent, agent, and user.

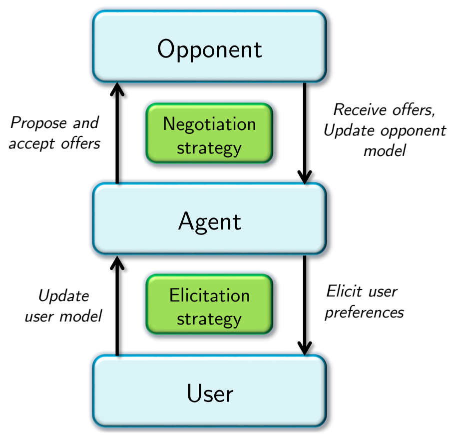

Figure 1: The interactions between the user, the agent, and the opponent.

To make the setting more concrete, consider the following scenario.

Example 1. *Suppose Alice has an automated agent purchasing software for her online. After some initial bargaining, the agent has received two proposals, ω_1 and ω_2, from the software supplier. There is another possible option ω_3, but there is only a 50% probability that the supplier will accept ω_3.*

The agent is not fully aware of Alice's preferences concerning each offer, but from previous negotiations, it deduces that: ω_1 is a risky choice, with utility anywhere between 0 and 1; ω_2 is a safer alternative, with utility between 0.4 and 0.6; ω_3 is the most certain and desirable choice, with utility between 0.8 and 0.9.

Not getting the software at all has utility 0.2, meaning ω_1 could turn out to be worse than opting out. Asking Alice for more information about ω_1, ω_2 or ω_3 lowers the utility of the outcome by 0.2 for every question. What is the agent to do?

Note that, although the expected utilities of ω_1 and ω_2 are the same, they clearly should be handled differently by the agent; for instance, if the agent asks Alice the true valuation of ω_1 and it turns out to be 0.9, the agent should immediately stop and settle for this option, while this does not hold for the other options.

To make this decision, the agent needs to optimally trade off the costs of user bother against the chances of getting the best deal for the user with the limited preference information available. We will return to this example later in the paper after presenting a formal framework to handle these cases.

3 Formal Model

Let $\Omega = \{\omega_1, \ldots, \omega_n\}$ denote the space of all possible outcomes of the negotiation. Every $\omega \in \Omega$ is a possible agreement, of which the utility for the user, denoted by $U(\omega)$, is *a priori* uncertain. Specifically, before any information is elicited from the user, the utility of an offer $\omega \in \Omega$ is given by a stochastic variable x_ω with cumulative distribution function $F_\omega(x)$, independent of the other offers.

At any point in time, the agent may ask the user to specify the actual utility $U(\omega)$ of a specific outcome ω at cost $c(\omega)$, which represents the utility loss associated with bothering the user. We assume that, once the user provides the utility, the information becomes certain. Note that the user need not provide this utility value directly: the preferences could be obtained by any elicitation method (e.g. queries that compare outcomes), as long as we can extract a reasonable utility value from this (e.g. using the Borda scoring rule as in [Aydoğan *et al.*, 2014]).

The agent may continue to elicit the user's preferences in an incremental fashion. In doing so, the user feedback partitions the outcome space into two sets: a growing set $S \subseteq \Omega$ of *known* offers, i.e. the set of offers for which the agent knows the utility for certain, and the set $\overline{S} = \Omega \backslash S$ of still-uncertain offers.

In addition, the agent models the probability that an offer $\omega \in \Omega$ is accepted by the opponent, which is denoted by p_ω. The prior distributions for p_ω can be constructed from earlier interactions and the opponent model is updated after each negotiation exchange so that it can take into account strategic behavior. This constitutes the agent's opponent model. Furthermore, the agent has a known reservation value $r \in [0, 1]$, which is the utility of a disagreement and hence the minimum acceptable utility for an offer by the user.

The agent exchanges offers with the opponent according to a negotiation protocol.[2] Whenever the agent receives an of-

[1]This information could be obtained from previous interactions or could be the population average. In this paper we are not concerned with how this information is obtained and assume it is available.

[2]In the experiments we use the alternating offers protocol, but

fer from the opponent, the opponent model is updated; similarly, the agent can propose an offer and updates the opponent model if it has been rejected. The negotiators can signal an accept by resending an offer that was proposed previously by the other. There is typically a non-zero probability the utility of an unknown offer falls below the reservation value; therefore, we allow the agent to only send and accept *known* offers, since otherwise the reservation value cannot be guaranteed. When an offer is accepted, the negotiation ends in agreement. An agreement must be reached before a fixed number of exchanges (or *rounds*) N. The agent may also choose to actively end the negotiation process by signaling a break-off. At the end of the negotiation, the utility of the agent is given by:

$$U = \begin{cases} U(\omega) - \sum_{\omega' \in S} c(\omega') & \text{if } \omega \in S \text{ is accepted,} \\ r - \sum_{\omega' \in S} c(\omega') & \text{if no agreement is reached.} \end{cases}$$

4 Negotiation and Elicitation Strategies

We now proceed to discuss the negotiation and elicitation strategies. Note that these strategies are tightly coupled since proposing the right offers depends on what is known about their utilities. Similarly, the elicitation strategy must consider the negotiation process when choosing which offers to elicit.

At any point in time, the agent needs to decide on its actions by comparing two quantities: 1) the utility the agent expects to obtain from the negotiation, given the set of currently known offers, called the *negotiation value*; and 2) the expected utility after one or more elicitation steps (i.e., selecting an x_ω and observing it at a cost). We define both notions in turn in Section 4.1 and 4.2.

4.1 The Negotiation Strategy

Given the set of known offers S from the elicitation strategy, at each round j of the negotiation, the agent needs to decide what offer to send to the opponent if any, whether to accept an incoming offer, or whether to break off the negotiation. In doing so, the agent's goal is to maximize expected utility. Therefore, we need to formulate the expected value of performing a certain action, taking into account the opponent model and any expected future utility if the offer is rejected and the negotiation proceeds to the next stage.

While it is, in principle, possible to reason about all possible future states of the negotiation and work backwards to define the negotiation value of an offer, this is computationally intractable. Instead, we assume that the agent uses a generic decision function which provides an *aspiration value*, $\alpha_j \in [0, 1]$, which depends on the current round $j \leq N$ and thus can change over the course of the negotiation. This allows our model to support a wide range of different negotiation strategies; e.g. conceding quickly, or playing a Boulware strategy, see also Section 5.1. The aspiration threshold represents 'what seems attainable' [Pruitt, 1981; Somefun *et al.*, 2004] and therefore acts as the expected reward of continuing the transaction when an offer gets declined. A popular approach of this sort is using *time-dependent tactics* [Faratin *et al.*, 1998; Fatima *et al.*, 2002]. However, the model does not restrict the types of functions that can be used, and the literature offers many other possibilities, e.g. [Chen *et al.*, 2013; Fatima *et al.*, 2002; Kawaguchi *et al.*, 2013].

Given the aspiration value α_j, we can formulate the *negotiation value* of sending an offer $\omega \in S$ as follows:

$$v(\omega) = p_\omega U(\omega) + (1 - p_\omega)\alpha_j,$$

which comprises its immediate payoff $U(\omega)$ if it gets accepted (with probability p_ω) and the expected future payoff α_j if it is rejected. To simplify notation, we augment S with a dummy offer ω_0 which has a known utility equal to the reservation value, i.e. $U(\omega_0) = r$.

Provided that there is no need to elicit any more offers at this point, then at each point in time during the negotiation process, the optimal strategy for the agent is simply to select the offer in S with the highest current negotiation value:

$$v^*(S) = \max_{\omega \in S} v(\omega). \tag{1}$$

Note that this approach also models accepting opponent offers and breaking off negotiations: if the chosen ω is the same as the offer which was sent by opponent, this signals an acceptance, while sending ω_0 signals a breakoff. This negotiation strategy is formalized in Algorithm 1.

Algorithm 1: A generic negotiation strategy.

Input: The current negotiation state.
Output: A break-off, a counter-offer or an accept.
begin
  for $\omega \in \Omega$ **do**
    `update`(p_ω);
  `callElicitationStrategy`();
  $\omega \longleftarrow \arg\max_{\omega' \in S} (p_{\omega'} U(\omega') + (1 - p_{\omega'})\alpha_j)$
  return $\begin{cases} \textsc{Breakoff} & \textit{if } \omega = \omega_0, \\ \textsc{Accept} & \textit{if } \omega \textit{ was offered}, \\ \textsc{Send}(\omega) & \textit{otherwise}. \end{cases}$

4.2 An Optimal Elicitation Strategy

Using the expected negotiation value $v^*(S)$ of a given set S of known offers, the elicitation strategy needs to determine which of the unknown offers, if any, to elicit from the user. The problem is non-trivial because the optimal elicitation strategy is sequential: whether or not to elicit an offer depends on the actual values of the offers elicited so far. Therefore, the goal is to find an optimal *sequence* of offers to elicit and a strategy which specifies when to stop the elicitation process.

Now, because the utilities in $\overline{S}$ are independently distributed, and the costs of elicited offers are sunk, it is easily verified that the elicitation strategy only depends on the set of currently unknown offers, $\overline{S}$, and the current negotiation value $v^*(S)$ that can be obtained. That is, the current *elicitation state* $\mathcal{E}$ can be summarized by $\langle \overline{S}, y \rangle$, which describes all non-elicited offers $\overline{S}$ plus the negotiation value $y = v^*(S)$ that is currently available.

---

the model is sufficiently generic to allow other protocols.

Our goal is to formulate an elicitation policy π, which chooses, given the state $\mathcal{E}$, whether to elicit another offer (if $\pi(\mathcal{E}) \in \overline{S}$), or to stop eliciting and proceed with the negotiation (if $\pi(\mathcal{E}) \notin \overline{S}$). The utility of a policy π can be computed as follows:

$$U(\pi, \mathcal{E}) = \begin{cases} y & \text{if } \pi(\mathcal{E}) \notin \overline{S}, \\ \int_{-\infty}^{\infty} U(\pi, \mathcal{E}') \, \mathrm{d}F_{x_{\pi(\mathcal{E})}}(x) & \text{otherwise,} \\ -c(\pi(\mathcal{E})) & \end{cases}$$

where $\mathcal{E}' = \langle \overline{S} \setminus \{\pi(\mathcal{E})\}, \max(y, v(x)) \rangle$ describes the new state after the stochastic variable $x_{\pi(\mathcal{E})}$ has been observed.

Given a state $\mathcal{E}$, we are looking for the optimal elicitation strategy $\pi^* = \arg\max_\pi U(\pi, \mathcal{E})$. Note that when all offers are known, $U(\pi^*, \langle \varnothing, y \rangle) = y$; otherwise, the agent may choose to elicit one or more offers $\omega \in \overline{S}$. Taking into account the negotiation process, the negotiation value of an unknown offer $\omega \in \overline{S}$ is given by the random variable:

$$x^v_\omega = p_\omega x_\omega + (1 - p_\omega)\alpha_j,$$

with $F^v_\omega(x)$ denoting the corresponding cumulative distribution function.

Given this, the optimal policy should consider the following: either stop the process and obtain y, or elicit an offer $\omega \in \overline{S}$ by sampling x from x^v_ω at cost $c(\omega)$, while taking the following into account:

- If $x \leq y$, nothing changes except the utility of ω is now known, and we have expected value $-c(\omega) + U(\pi^*, \langle \overline{S} \setminus \{\omega\}, y \rangle)$;
- Otherwise, we have found a better reward $x > y$, and the expected utility is: $-c(\omega) + U(\pi^*, \langle \overline{S} \setminus \{\omega\}, x \rangle)$.

More formally, $U(\pi^*, \mathcal{E})$ must satisfy the following recursive relation:

$$U(\pi^*, \mathcal{E}) = \max \left\{ y, \max_{\omega \in \overline{S}} \left\{ U(\pi^*, \langle \overline{S} \setminus \{\omega\}, y \rangle) \cdot F^v_\omega(y) \right. \right. \tag{2}$$
$$\left. \left. -c(\omega) + \int_{x=y}^{\infty} U(\pi^*, \langle \overline{S} \setminus \{\omega\}, x \rangle) \, \mathrm{d}F^v_\omega(x) \right\} \right\}.$$

The relation for π^* given in Eq. (2) is essentially a Bellman equation which, in principle, could be solved by backward induction. However, even for a moderate-size negotiation space, this approach quickly becomes intractable. We provide a simple (but optimal) index-based method to decide, in $O(n \log n)$ time, which offers to elicit from the user and to determine whether to stop eliciting. This index, denoted z^v_ω, provides us with an alternative to solving Eq. (2), and is defined as the solution to:

$$\int_{z^v_\omega}^{\infty} (x - z^v_\omega) \, \mathrm{d}F^v_\omega(x) = c(\omega). \tag{3}$$

Given state $\mathcal{E} = \langle \overline{S}, y \rangle$, we are now ready to formulate our elicitation strategy:

Elicitation Strategy π^***.** *Elicit the unknown offer with the highest index z^v_ω if it is higher than $v^*(S)$; update the $v^*(S)$ if the realized value is higher and repeat the process. Stop the elicitation process as soon as the highest index is less than $v^*(S)$, or when all offers are known.*

Proposition 4.1. *The elicitation strategy π^* maximizes expected reward $U(\pi^*, \mathcal{E})$.*

Proof. We show that our formulation of the problem can be mapped to Pandora's problem [Weitzman, 1979], which is a question in search theory about opening boxes. Specifically, each offer in Ω can be viewed as a box, where S (including ω_0) is the set of open boxes with reward of $p_\omega U(\omega) + (1 - p_\omega)\alpha_j$. Unknown offers $\omega \in \overline{S}$ are closed boxes with distributions $F^v_\omega(x)$ and cost $c(\omega)$, and with associated index z^v_ω. Furthermore, the reward of the open boxes S is given by the negotiation value $v^*(S)$ (Eq. 1). Weitzman [1979] shows that, for a closed box with a reward distributed according to the cumulative distribution function F and with opening cost c, we can assign an index z, satisfying $\int_z^\infty (x - z) \, \mathrm{d}F(x) = c$ which fully captures the relevant information about the closed box: it should be opened exactly when it has the highest index and exceeds the reward in one of the open boxes. It is proven in [Weitzman, 1979] that this strategy is optimal in terms of expected reward (Eq. 2). □

The optimal elicitation strategy is shown in Algorithm 2.

Algorithm 2: Using Pandora's Rule to formulate an optimal elicitation strategy.

```
begin
    for ω ∈ S̄ do
        z^v_ω ⟵ Solve ∫_z^∞ (x − z) dF^v_ω(x) = c(ω) for z;
    v ⟵ max_{ω∈S}(p_ω U(ω) + (1 − p_ω)α_j)
    Loop
        ω ⟵ arg max_{ω'∈S̄} z^v_{ω'};
        if z^v_ω < v or S̄ = ∅ then
            return;
        else
            U(ω) ⟵ elicitFromUser(ω);
            elicitationCost ⟵ elicitationCost + c(ω);
            S ⟵ S ∪ {ω}; S̄ ⟵ S̄ \ {ω};
            v ⟵ max (v, p_ω U(ω) + (1 − p_ω)α_j);
```

Proposition 4.2. *The complexity of the elicitation strategy is $O(n \log n)$.*

Proof. Note that, crucially, the index values do not need updating after Algorithm 2 has chosen to elicit a particular offer; hence, to find the offer in $\overline{S}$ with the highest index, it suffices to order the set of indexes $\{z^v_\omega \mid \omega \in \overline{S}\}$ once, in $O(n \log n)$ time, before entering the loop in Algorithm 2. □

Our algorithm has a number of desirable properties. Firstly, if it is costless to acquire information about an offer ω, then the algorithm will always elicit it, provided it is not dominated by an offer already in S:

Proposition 4.3. *Let $\omega \in \overline{S}$ be such that $c(\omega) = 0$ and $F^v_\omega(v^*(S)) < 1$. Then ω will be elicited by Algorithm 2.*

Proof. Olszewski and Weber [2012] show that, for an index z^v_ω satisfying Eq. (3), we have:

$$z^v_\omega = \min\{y \mid E\left[\max\left(0, x^v_\omega - y\right)\right] \leq c(\omega) \wedge y \geq 0\}.$$

It follows that, if $c(\omega) = 0$, then z^v_ω will be the supremum for the range of x^v_ω. Hence, $z^v_\omega \geq v^*(S)$, which guarantees ω gets elicited in Algorithm 2. □

A second property is the negotiation behavior induced by Algorithm 2: if the aspiration α_j exceeds the utility of the known offers, the algorithm tends to select risky offers; i.e., high utility offers with a low probability of acceptance. This follows from the fact that, in this case, $p_\omega U(\omega) + (1 - p_\omega)\alpha_j$ is maximized for low p_ω and high $U(\omega)$. Conversely, when α_j approaches the reservation value r, the algorithm will optimize immediate reward, which is especially useful towards the end of the negotiation. The agent will also always seek to accept any received offer higher than the reservation value rather than breaking off, as $p_\omega U(\omega) + (1 - p_\omega)\alpha_j > r$ for any $U(\omega) > r$.

To further illustrate the behavior of the algorithm, we conclude this section by providing a solution to our example of Alice's software purchasing agent.

Example 1 (continued)**.** *Let us assume that the probability functions of all options are distributed uniformly, and that the negotiation is drawing to a close, so that $\alpha_j = r = 0.2$. Using Eq. (3), the index of ω_1 can be computed to be $z^v_{\omega_1} = \frac{1}{5}(5 - \sqrt{10}) \approx 0.37$, while the index of ω_2 is $z^v_{\omega_2} = 0.3$. For ω_3, we have $p_{\omega_3} = 0.5$, so*

$$x^v_{\omega_3} = \frac{1}{2}U(0.8; 0.9) + \frac{1}{2} \cdot 0.2 = U(0.5; 0.55),$$

and from that, we can compute $z^v_{\omega_3} = \frac{13}{40} = 0.325$.

This means that the optimal elicitation strategy is as follows: since $z^v_{\omega_1} > z^v_{\omega_3} > z^v_{\omega_2}$, the agent should first extract Alice's utility value for the riskier option ω_1. If $U(\omega_1)$ turns out to be higher than $z^v_{\omega_3} = 0.325$, it should be accepted; otherwise, ω_3 should also be elicited, and the agent should send out this offer regardless of the outcome, as $\frac{1}{2}U(\omega_3) + 0.1 > 0.3$.

5 Experiments

To analyze the performance of our elicitation strategy, we test it across a wide range of negotiation scenarios against a set of benchmark elicitation strategies. All elicitation strategies are tested on the same set of scenarios, with the same set of aspiration thresholds, and all of them use Algorithm 1 to decide on their offer; they only differ in the way they elicit offers. This means differences in results can be attributed exclusively to the effectiveness of each elicitation strategy.

5.1 Setup

For our experiments, the agent exchanges bids with the opponent using the alternating offers protocol [Osborne and Rubinstein, 1994] with a deadline of $N = 10$ and $N = 100$ rounds. In the initial state of every negotiation, $S = \varnothing$ and $\overline{S} = \Omega$, and hence the agent's starting information about the user's preferences is limited to the stochastic distributions x_ω.

We select 200 different negotiation scenarios with $|\Omega| = 10$. In each scenario, we select a different utility probability function F_ω for every $\omega \in \Omega$ by setting it to either a uniform distribution $U(a, b)$, with $a < b$ uniformly sampled from $U(0, 1)$, or a beta distribution $Beta(\alpha, \beta)$, with $\alpha, \beta \in \{1, \ldots, 10\}$. To determine $U(\omega)$, we take a random sample from x_ω; in other words, x_ω is an accurate stochastic representation of $U(\omega)$. The reservation value r acts as a reference value, which we set to 0.25. We vary the elicitation costs $c(\omega)$ as a parameter of the experiments, choosing values between 0 and 1 with 0.05 increments.

For every $\omega \in \Omega$, we initialize the agent's opponent model by sampling p_ω from $U(0, 1)$. The learning method of the agent is elementary: p_ω retains its initial value, unless ω is rejected by the opponent, in which case it is set to $p_\omega = 0$. Conversely, if ω is offered by the opponent, we set $p_\omega = 1$.

The opponent has a fixed set of offers B (unbeknown to the agent) that it will accept, and from which it picks randomly when proposing offers (or ending the negotiation in case $B = \varnothing$). Every offer $\omega \in \Omega$ has a probability of p_ω to be included in the set of acceptable offers B.

As every opponent acts in a stochastic manner, we repeat every negotiation in all of the 200 scenarios 10 times to increase statistical significance of our results.

For setting the aspiration threshold α_j of the agent, we use the myopic version of our algorithm that maximizes immediate payoff in the current round ($\alpha_j = 0$) as well as well-known *time-dependent tactics* [Faratin *et al.*, 1998; Fatima *et al.*, 2002]. Specifically, in round $j \leq N$ of the negotiation, this family of strategies aims for utility closest to:

$$\alpha_j = P_{\text{min}} + (P_{\text{max}} - P_{\text{min}}) \cdot (1 - (j/N)^{1/e}).$$

We set $P_{\text{min}} = r$ so that the aspiration threshold reaches the reservation value at the deadline, and we set P_{max} to $1/2$, which is a reasonable choice for the majority of elicitation costs. Lastly, we select three types of time-dependent tactics to define the aspiration threshold: *Boulware* ($e = 5$), *Linear* ($e = 1$), and *Conceder* ($e = \frac{1}{5}$).

5.2 Benchmarks

We compare our elicitation strategy with four benchmark strategies. Two of them are included as baseline strategies: first, we consider a policy in which the agent elicits an offer at *random* at every negotiation round. Our second baseline strategy elicits *all* offers, which is worthwhile for small user bother costs.

Next, a reasonable benchmark strategy is to elicit the offers with the *highest expected value minus the cost*; i.e., $\max_{\omega \in \overline{S}} E(x^v_\omega) - c(\omega)$. In a sense, this is a first-order approximation of Pandora's Rule, as Eq. (3) simplifies to this for sufficiently large elicitation cost. We also consider a variant of this strategy which, in addition, elicits all *incoming* offers.

Finally, we include a theoretical upper bound that represents the maximum utility that can be obtained by an elicitation strategy with perfect foresight; i.e., the agent pays elicitation costs, but is allowed to cheat by using a perfect opponent model.

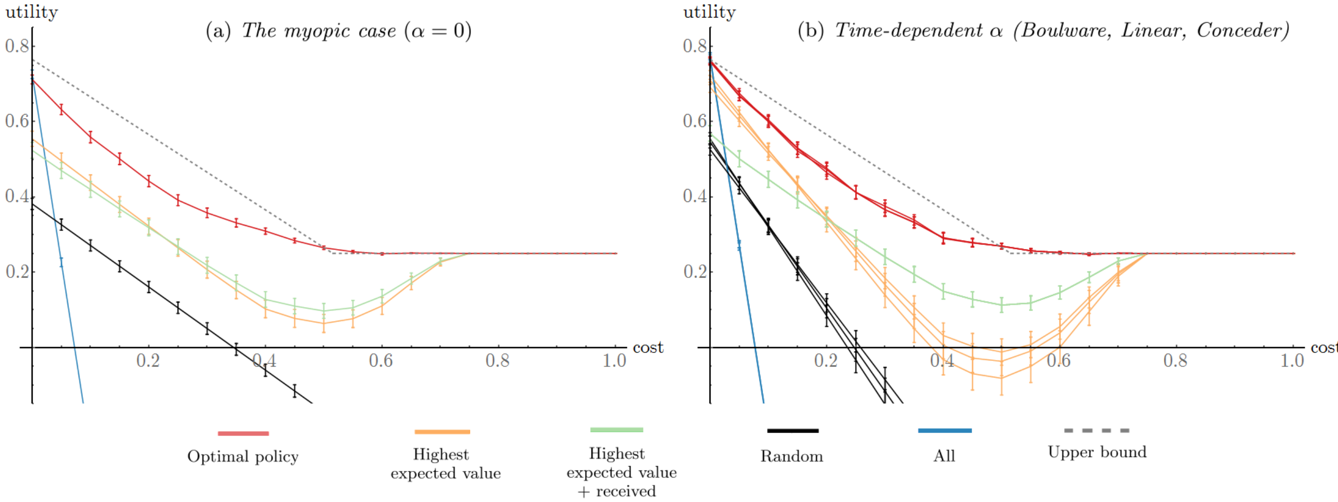

Figure 2: The performance of five different elicitation techniques for elicitation costs between 0 and 1. Error bars indicate one standard error difference to the mean.

5.3 Results

Figure 2 shows the average utility obtained by every elicitation strategy in our experiments, as well as the theoretical upper bound, for varying elicitation costs. Figure 2a shows the results for $\alpha_j = 0$, while Figure 2b shows the superimposed results of three different time-dependent aspiration thresholds. The results are given for $N = 10$ rounds and uniform utility distributions; the results are very similar for $N = 100$ rounds and for beta utility distributions, but we omit them due to space constraints. The standard error is measured over the different scenarios.

As is evident from Figures 2a and 2b, our optimal policy significantly outperforms all others for all aspiration thresholds (t-test, $p < 0.05$), acquiring utility higher than the other methods across all elicitation costs in the crucial interval $[0.1, 0.7]$. As expected, the optimal strategy's payoff is higher for lower costs, and peaks at zero costs, exactly overlapping the utility obtained by eliciting all offers. This confirms our earlier statement that our algorithm should work optimally in this case (see Section 4.2). The obtained utility slowly declines until around 0.65, where the bother costs are too high to conduct a meaningful negotiation. In this case, it is optimal to not elicit any offers and instead to break off the negotiation, thereby earning $r = 0.25$. The performance of our optimal policy comes surprisingly close[3] to the theoretical upper bound, considering that the latter uses perfect information.

The strategy that elicits all offers performs well for low elicitation costs; however, for higher costs, the performance of this strategy quickly degenerates and moves off the chart, because of the rapid increase of the total costs $\sum_{\omega\in\Omega} c(\omega)$. For higher elicitation costs, the random elicitation method works slightly better, but performs poorly overall. In most cases, the aspiration threshold set by *Boulware* performs slightly worse than *Linear*, which in turn is outperformed by *Conceder*. The differences are small though, especially for the optimal policy, which indicates its robustness.

The benchmark strategies that elicit offers with the highest expected reward (both with and without accepting incoming offers) together count as second best. Eliciting incoming offers is a safer choice for higher elicitation costs, because such offers are sure to be accepted by the opponent. The difference with the optimal policy's performance is significant, and stems from the fact that our optimal policy takes future reward into account when exploring the different options.

In general, the optimal policy will initially elicit the high-risk/high-reward offers, as it knows when to incrementally elicit other options in case this does not pay off. This look-ahead behavior is able to balance exploration and exploitation in an optimal way.

6 Discussion and Future Work

In this paper, we deal with the problem of representing a user in a negotiation with only limited preference information available. Our optimal elicitation strategy is efficient and is generally applicable to a wide range of bidding strategies. Our results indicate that our method performs well under a variety of circumstances and significantly outperforms other benchmark strategies in a robust way. By directly incorporating the user into the decision making process, our method can be used to provide negotiation support in a variety of complex negotiation domains, ranging from purchasing a supercomputer to conducting union negotiations.

In our model we make no assumptions about the structure of the utility function $U(\omega)$, which could be drawn from offline aggregate data and techniques such as collaborative filtering. However, in some particular cases, for example for linear additive utility functions, it would be possible to elicit the user's preferences more effectively by obtaining information about the issue weights instead. In future work, it would be interesting to apply Pandora's Rule to such a setting, as

[3] Note that, although the upper bound seems to momentarily fall below the graph of our optimal policy, none of the sampled points actually exceed the upper bound; this is purely an artifact of linear interpolation between the samples.

this would require a separate decision layer that assigns the *value of information* to each elicitation action, in the spirit of the work by Chajewska et al. [1998].

Another general feature of our approach is that we allow for different costs for each elicitation action. This means our method works for any user bother cost function, which, as a possible next step, could be determined by methods from HCI and elicitation theory. This way, we could differentiate, for example, between more patient and impatient users, and take into account how much recent bother the user has been subjected to, along the lines of [Buffett *et al.*, 2004].

Acknowledgments

This work was supported by the EPSRC, grant number EP/K039989/1.

References

[Aydoğan *et al.*, 2014] Reyhan Aydoğan, Tim Baarslag, Koen V. Hindriks, Catholijn M. Jonker, and Pınar Yolum. Heuristics for using CP-nets in utility-based negotiation without knowing utilities. *Knowledge and Information Systems*, pages 1–32, 2014.

[Boutilier, 2002] Craig Boutilier. A POMDP formulation of preference elicitation problems. In *Eighteenth National Conference on Artificial Intelligence*, pages 239–246, Menlo Park, CA, USA, 2002. American Association for Artificial Intelligence.

[Buffett *et al.*, 2004] Scott Buffett, Nathan Scott, Bruce Spencer, Michael Richter, and Michael W. Fleming. Determining internet users' values for private information. In *Second Annual Conference on Privacy, Security and Trust, October 13-15, 2004, Wu Centre, University of New Brunswick, Fredericton, New Brunswick, Canada, Proceedings*, pages 79–88, 2004.

[Chajewska *et al.*, 1998] Urszula Chajewska, Lise Getoor, Joseph Norman, and Yuval Shahar. Utility elicitation as a classification problem. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, UAI'98, pages 79–88, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers Inc.

[Chajewska *et al.*, 2000] Urszula Chajewska, Daphne Koller, and Ronald Parr. Making rational decisions using adaptive utility elicitation. In *In Proceedings of the Seventeenth National Conference on Artificial Intelligence*, pages 363–369, 2000.

[Chen *et al.*, 2013] Siqi Chen, Haitham Bou Ammar, Karl Tuyls, and Gerhard Weiss. Optimizing complex automated negotiation using sparse pseudo-input gaussian processes. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pages 707–714, Richland, SC, 2013. International Foundation for Autonomous Agents and Multiagent Systems.

[Conen and Sandholm, 2001] Wolfram Conen and Tuomas Sandholm. Minimal preference elicitation in combinatorial auctions. In *In Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, Workshop on Economic Agents, Models, and Mechanisms*, pages 71–80, 2001.

[Faratin *et al.*, 1998] Peyman Faratin, Carles Sierra, and Nick R. Jennings. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, 24(3-4):159 – 182, 1998. Multi-Agent Rationality.

[Fatima *et al.*, 2002] Shaheen S. Fatima, Michael Wooldridge, and Nicholas R. Jennings. Multi-issue negotiation under time constraints. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 143–150, New York, NY, USA, 2002. ACM.

[Fatima *et al.*, 2014] Shaheen Fatima, Sarit Kraus, and Michael Wooldridge. *Principles of Automated Negotiation*. Cambridge University Press, 2014.

[Fleming and Cohen, 2004] Michael Fleming and Robin Cohen. A decision procedure for autonomous agents to reason about interaction with humans. In *Proceedings of the AAAI 2004 Spring Symposium on Interaction between Humans and Autonomous Systems over Extended Operation*, pages 81–86, 2004.

[Kawaguchi *et al.*, 2013] Shogo Kawaguchi, Katsuhide Fujita, and Takayuki Ito. AgentK2: Compromising strategy based on estimated maximum utility for automated negotiating agents. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, pages 235–241. Springer Berlin Heidelberg, 2013.

[Larson and Sandholm, 2001] Kate Larson and Tuomas Sandholm. Costly valuation computation in auctions. In *Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '01, pages 169–182, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[Lomuscio *et al.*, 2001] Alessio Lomuscio, Michael Wooldridge, and Nicholas Jennings. A classification scheme for negotiation in electronic commerce. In Frank Dignum and Carles Sierra, editors, *Agent Mediated Electronic Commerce*, volume 1991 of *Lecture Notes in Computer Science*, pages 19–33. Springer Berlin Heidelberg, 2001.

[Olszewski and Weber, 2012] Wojciech Olszewski and Richard Weber. A more general pandora rule. Technical report, Mimeo, 2012.

[Osborne and Rubinstein, 1994] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*, volume 1. The MIT Press, 1st edition, 1994.

[Parkes, 2005] David C. Parkes. Auction design with costly preference elicitation. *Annals of Mathematics and Artificial Intelligence*, 44(3):269–302, 2005.

[Pruitt, 1981] D. G. Pruitt. *Negotiation Behavior*. Academic Press, 1981.

[Somefun *et al.*, 2004] D.J.A. Somefun, E.H. Gerding, S. Bohte, and J.A. La Poutré. Automated negotiation and bundling of information goods. In Peyman Faratin, David C. Parkes, Juan A. Rodríguez-Aguilar, and William E. Walsh, editors, *Agent-Mediated Electronic Commerce V. Designing Mechanisms and Systems*, volume 3048 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin Heidelberg, 2004.

[Weitzman, 1979] Martin L Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.