Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence

Negotiation in Exploration-Based Environment

Israel Sofer and **David Sarne** and **Avinatan Hassidim**
Department of Computer Science
Bar Ilan University, Ramat Gan 52900, Israel

Abstract

This paper studies repetitive negotiation over the execution of an exploration process between two self-interested, fully rational agents in a full information environment with side payments. A key aspect of the protocol is that the exploration's execution may interleave with the negotiation itself, inflicting some degradation on the exploration's flexibility. The advantage of this form of negotiation is in enabling the agents supervising that the exploration's execution takes place in its agreed form as negotiated. We show that in many cases, much of the computational complexity of the new protocol can be eliminated by solving an alternative negotiation scheme according to which the parties first negotiate the exploration terms as a whole and then execute it. As demonstrated in the paper, the solution characteristics of the new protocol are somehow different from those of legacy negotiation protocols where the execution of the agreement reached through the negotiation is completely separated from the negotiation process. Furthermore, if the agents are given the option to control some of the negotiation protocol parameters, the resulting exploration may be suboptimal. In particular we show that the increase in an agent's expected utility in such cases is unbounded and so is the resulting decrease in the social welfare. Surprisingly, we show that further increasing one of the agents' level of control in some of the negotiation parameters enables bounding the resulting decrease in the social welfare.

Introduction

Negotiation is one of the most explored fields in artificial intelligence and a key concept in multi-agent systems. Through the process of negotiation, agents can decide how to split gains achieved from cooperation (Lomuscio, Wooldridge, and Jennings 2001; Jennings et al. 2001; Rosenschein and Zlotkin 1994). The simplest form of negotiation involves two agents, where in each step of the negotiation one of the agents makes an offer, and the other agent decides whether to accept or reject it. The agents usually take turns in being the agent who makes the offer. The outcome of the negotiation depends on the agents strategies which, in turn, depend on the protocol and the information that the agents have about the negotiation parameters, such as: their reserve prices, the utility functions and private deadlines. The information the agents have can be full, incomplete or change during the negotiation. The negotiation can be over a single issue, over multiple issues that may be negotiated sequentially (i.e., only one issue in each negotiation step) or over a bundle of issues. All of these concepts generalize to protocols in which more than two agents participate in the negotiation.

This paper studies a negotiation protocol with two fully rational agents who have complete information on each other's preferences. The agents are faced with a set of opportunities, which values can be revealed only through costly exploration. The situation is further complicated by the fact that only one of the agents can carry out the exploration, whereas only the other agent benefits from the information revealed through it. Such a setting characterizes many real-life applications. For example, consider the case of oil drilling, where one side is a contractor, capable of drilling, and the other side is the company that purchased the drilling rights for the area. The company pays the contractor for drilling and benefits from the findings. Here, there is a cost for every drill and gains change according to the amount of oil found during that drilling. Another example is technological research and development. The tech company may request that an external research lab (or a professor receiving a grant) explore several avenues in order to choose the most promising one. The research lab has to pay the cost (do the exploration work) for every such avenue.

A key aspect in our setting is that the agents do not decide beforehand how the exploration will be performed and how much the beneficiary agent will pay the exploring agent, but rather negotiate it as the exploration progresses. Thus, the outcomes of the negotiation depend on the results obtained from the exploration process as it progresses. Similarly, the agents negotiate not only over the amount that the beneficiary agent will pay the exploring agent, but also over what opportunities to explore and when. To model this, at each step of the negotiation one of the agents makes a proposal to the other agent which includes what opportunity will be explored next and how much the beneficiary agent will pay the exploring agent for doing so. Then the other agent decides whether to accept the proposal or not. An acceptance is followed by exploring the appropriate opportunity and the negotiation continues after. One key advantage of this interleaved negotiation and exploration process (herein denoted "interleaved protocol") is that the agent benefiting from the exploration can make sure that the other agent executes the exploration as agreed upon. Furthermore, this form of negotiation is often more intuitive for people as it does not require agreeing on a complex exploration strategy in advance.

One important aspect that arises when considering the interleaved protocol is its complexity. In the traditional non-interleaved protocols, where the agents first negotiate and then act (Rubinstein 1982), computing the offers made by the agents can be done by starting from the deadline, and working our way backwards to the first offer. A key requirement for this backward induction is that no new inputs enter the negotiation, and thus in the first step one can predict the entire behavior. When the negotiation is interleaved with exploration, this basic assumption no longer holds, as every exploration step changes the situation of the agents (they learn the value of one of the opportunities). Thus, in order to make the first offer, the agent has to consider all possible outcomes of the exploration steps, and the complexity becomes exponential in the number of opportunities.

The first part of this paper shows equivalence between the setting in which negotiation and exploration are interleaved with a simpler setting in which negotiation and exploration are separated, and the exploration is carried out as a whole under a global time constraint (which applies both to the negotiation and exploration) upon reaching an agreement. The equivalent setting has the same expected utilities, and enables efficient computation of the offers made in the setting in which exploration and negotiation are interleaved. The complexity of the new algorithm is the same order of magnitude as solving the exploration problem (without the negotiation).

We then use the equivalent setting, to study properties of the interleaved setting. We show that the agents' utilities are influenced by the discounting factor and cost and by the negotiation horizon, and demonstrate the balance between them. We generalize this to allow the exploration of several opportunities at once, and show that the tradeoffs are different from the ones that occur in settings where the agents are concerned with just the exploration problem (as a stand-alone problem).

Finally, we study the case where one agent has more control over the negotiations. In one setting, we allow one of the agents to choose the number of opportunities which will be explored every time (this number is fixed and the proposing agent gets to choose what opportunities to propose). We show that in this case the ratio between the agent utility with the extra control and its utility in the social welfare maximizing case (in the symmetric setting) can be unbounded. Further, we show that the optimal strategy for that agent (which maximizes its utility) can greatly decrease the social welfare (almost to zero), compared to the symmetric case. In another setting, we let one of the agents decide how many opportunities will be negotiated at each step. Although this setting may look similar to the previous one, it differs in the sense that the decrease in social welfare has a bound.

Related Work

Multi-agent negotiation is an active research area. The problem is defined by the negotiation space (also called negotiation protocol), which typically includes a negotiation protocol (i.e the set of the interaction rules between the agents), negotiation objectives (i.e. the range of issues to negotiate on) and negotiation strategies (which are the sequences of actions that the agents plan to take in order to achieve their objectives) (Fershtman 1990; Jennings et al. 2001). The agents must reach an agreement about the negotiation protocol before the negotiation begins and the strategy of the agents might change according to the negotiation protocol (e.g, (Lomuscio, Wooldridge, and Jennings 2001; Rosenschein and Zlotkin 1994)). See (Buttner 2006) for a survey on the topic.

The analysis found in the negotiation literature encompasses various protocols, differing in the assumptions they make regarding the negotiation mechanism, mostly in terms of the information negotiators have about their environment (Rubinstein 1982) the other negotiators types (Rubinstein 1985), the negotiation deadline, the level of rationality of the negotiators (ranging from fully rational (Sandholm and Vulkan 1999; Faratin, Sierra, and Jennings 2002) to bounded rational (Lin et al. 2008)), their level of cooperation (self-interested versus fully cooperative agents (Kraus 1997)) and the number of issues negotiated (either a single issue or a number of issues (Fatima, Wooldridge, and Jennings 2004; Lin and Chou 2004)). The negotiation protocol often includes a discounting factor and a negotiation deadline that influence the negotiation strategy (Ma and Manove 1993).

The main difference between the above related literature and the protocol presented in this paper is that prior work usually assumes that negotiation and execution are separate processes. That is, once an agreement is reached, the negotiation terminates and the execution begins. In our protocol this is not the case: the agents reach an agreement, execute it (which gives them new information), and then go back to the negotiation table to discuss the next agreement. The agents' strategies in the new negotiation can depend on the results of the execution of the previous agreements. This substantially complicates the process of computing the agents' strategies, as discussed in the analysis section. Some of the related work presents protocols where computational complexity is an issue (Dunne, Wooldridge, and Laurence 2005; Chevaleyre, Endriss, and Maudet 2010). Nevertheless, the increased complexity in these works is the result of the optimization problem (e.g., the optimization problem NP-complete (Larson and Sandholm 2002)) rather than the effect of continuing the negotiation process after an agreement is reached and executed. Works considering multi-issue negotiation do consider the situation where the negotiation resumes upon reaching an agreement (Fatima, Wooldridge, and Jennings 2006; Bac and Raff 1996), however they assume that the actual execution of the agreements is carried out after all agreements are made. Therefore, the uncertainty of executions does not play any role in these protocols. Moreover, work in human computer decision making has proposed agent designs that interleave repeated negotiation with execution but do not employ exploration strategies (Gal et al. 2011).

The costly exploration problem embedded in our negotiation protocol is standard (McMillan and Rothschild 1994; Chhabra, Das, and Sarne 2011). Over the years, many optimal exploration strategies have been introduced to various variants of the model (Stigler 1961; Weitzman 1979). Nevertheless, despite considering settings where agents cooperate in exploration (Sarne, Manisterski, and Kraus 2010), the optimal exploration literature has not addressed exploration as part of a negotiation setting.

The Negotiation Protocol

The protocol assumes a negotiation setting with two agents, Agt_1 and Agt_2 facing a costly exploration process, where Agt_1 is the one benefiting from the exploration and Agt_2 is the one actually capable of carrying out the exploration. Both agents are assumed to be fully rational and self-interested, thus Agt_2 will perform the exploration only for an appropriate compensation. The exploration setting consists of a set of n opportunities $B = \{b_1, \ldots, b_n\}$. Each opportunity b_i encapsulates a value, v_i unknown to the agents. The agents are acquainted, however, with the probability density function from which an opportunity b_i's value derives, denoted $f_i(x)$. This value can be obtained through exploration of the opportunity, incurring a fee (cost), denoted c_i, possibly different for each opportunity. While any explored opportunity is applicable for Agt_1, it is capable of exploiting only one. Thus, given several opportunities in which values were revealed, the agent prefers exploiting the highest one.

The negotiation protocol consists of sequential interleaved negotiation and exploration steps: on each negotiation step the proposing agent offers the exploration of a single opportunity b_i (by Agt_2) in exchange for a payment M (made by Agt_1 to Agt_2). If an agreement is reached, Agt_2 explores the opportunity and the agents advance to the next negotiation step. The negotiation takes place according to an alternating-offers protocol, by which the former proposer becomes the responder and vice-versa at each negotiation step. The utility of Agt_2 out of the negotiation, denoted U_2, is the discounted sum of payments it receives throughout the negotiation and the incurred exploration fees. The utility of Agt_1, denoted U_1, is the maximum among the discounted values of the opportunities revealed by the time the negotiation terminates minus the discounted sum of payments it makes to Agt_2. The negotiation is limited to a maximum of T negotiation steps. A similar discounting factor δ is used for gains, explorations costs and payments made. In the interest of simplicity the protocol assumes the cost incurred by the exploration of an opportunity is paid at the beginning of the exploration and the value of the opportunity is obtained at the beginning of the next negotiation round.

Analysis

We first review the solution to the exploration problem if carried out stand-alone and then analyze the solution to the negotiation problem. The exploration setting embedded in the interleaved negotiation protocol can be mapped to the canonical sequential exploration problem described by (Weitzman 1979). The optimal exploration strategy in this case is the one that maximizes the expected value obtained when the process terminates minus the expected sum of the costs incurred along the exploration. The complexity of solving the stand-alone exploration problem is setting-dependent. For example, if the opportunities are homogeneous (i.e., associated with the same distribution of values and exploration costs) then the solution is threshold-based and can be calculated in a rather straightforward manner (McMillan and Rothschild 1994). So is the case when the opportunities are heterogeneous, however the agent can potentially explore them all if it requested to do so (Weitzman 1979). In other cases, e.g., when there are heterogeneous opportunities and a bound on the number of opportunities which can be explored (the exploration horizon is shorter than the number of opportunities) the choice of the next opportunity to explore may depend on other factors, and the computational complexity may increase substantially.

The analysis of the negotiation protocol uses a standard backward induction technique (Stahl 1972). Each agent, when acting as the proposer, offers the exploration of a specific opportunity and a payment that guarantees that the responding agent is indifferent to accepting or rejecting the proposal (and thus the offer is necessarily accepted). We use $U_P(t, v, \mathcal{B})$ to denote the expected utility gain (onwards) of agent P if reaching negotiation step t with a set $\mathcal{B} = \{b_{i_1}, \ldots, b_{i_m}\} \subset B$ of opportunities that can still be explored and the best known value is v. In the case of Agt_1 $U_P(t, v, \mathcal{B})$ captures the discounted expected improvement in its exploitation value due to the explorations to come minus the expected discounted payments made to Agt_2 from this step onwards. Similarly, for Agt_2 $U_P(t, v, \mathcal{B})$ represents the expected discounted payments received minus the exploration costs incurred throughout future explorations.

Table 1 presents the value of $U_P(t, v, \mathcal{B})$ for each of the players for the cases where the proposer decides to terminate the negotiation and when a proposal to explore opportunity b_{i_j} for a payment $M(t, v, \mathcal{B})$ is made and is either accepted or rejected. Notice that, because the process necessarily terminates after the $T-th$ negotiation round, $U_P(T+1, v, \mathcal{B}) = 0$.

It is notable that for any negotiation state $(t, v, \mathcal{B})$, the sum of the utilities $U_1(t, v, \mathcal{B}) + U_2(t, v, \mathcal{B})$, denoted $EV(t, v, \mathcal{B})$, equals the expected improvement in Agt_1's exploitation value minus the costs that Agt_2 incurs throughout the exploration that takes place afterwards. This can be explained by the fact that, according to Table 1, the only changes in the agent utilities that are not associated with internal payments between the agents are in fact those two elements. The value of $EV(t, v, \mathcal{B})$ thus equals the expected utility of a single agent facing the stand-alone exploration problem characterized as $(t, v, \mathcal{B})$ if following the same exploration that the negotiation yields.

At any step of the negotiation, the proposer will choose either to terminate the negotiation or to make an offer in a way that its expected utility gain is maximized. A proposal will be accepted only if it guarantees the responder at least the same expected utility gain it achieves from rejecting it. The payment $M(t, v, \mathcal{B})$ suggested/requested as part of a proposal made by the proposer at any step of the negotiation t for exploring opportunity b_{i_j} is thus the discounted difference between the responder's utilities when accepting and rejecting the proposal, according to Table 1. Such a proposal will be made only if the utility to the proposer from

	$U_1(t,v,\mathcal{B})$	$U_2(t,v,\mathcal{B})$
Negotiation terminates	0	0
Proposal accepted	$-M(t,v,\mathcal{B})+\delta\int_{y=-\infty}^{\infty} max(v,y)f_i(y)dy - v$ $+\delta\int_{y=-\infty}^{\infty} U_1(t+1, max(v,y), \mathcal{B}-b_{i_j})f_i(y)dy$	$\delta\int_{y=-\infty}^{\infty} U_2(t+1, max(v,y), \mathcal{B}-b_{i_j})f_i(y)dy$ $+M(t,v,\mathcal{B})-c_i$
Proposal rejected	$\delta U_1(t+1,v,\mathcal{B})$	$\delta U_2(t+1,v,\mathcal{B})$

Table 1: The expected utility gain (onwards) of agent P if reaching negotiation step t with a set $\mathcal{B}=\{b_{i_1},\ldots,b_{i_m}\}\subset B$ of opportunities that can still be explored and the best known value is v.

its acceptance by the responder yields a positive expected utility.

The new negotiation protocol is more complex than the classic non-interleaved negotiation protocol (Rubinstein 1982), where the agents first negotiate the exploration as a whole and the agreed exploration is executed afterwards (with no time constraints over the exploration process). In the latter form of negotiation the computational complexity of calculating the solution is linear in the number of negotiation steps, since the negotiation terminates once an offer is received. In the interleaved protocol the complexity is exponential since regardless whether an offer is accepted or not the negotiation may continue. In order to deal with the computational complexity we introduce Theorem 1.

Theorem 1. *Given the current negotiation step t, a subset of unexplored opportunities, $\mathcal{B}$, and the current exploitation value, v, the agents' expected utilities, $U_1(t,v,\mathcal{B})$ and $U_2(t,v,\mathcal{B})$ resulting from negotiating the exploration of opportunities one at a time (and executing the exploration upon each agreement) are equal to those resulting from continuing the negotiation in a way that the exploration is first negotiated as a whole (without limiting proposals to a single exploration) and only then executing the exploration within the limits of the remaining steps.*

Proof. *We first prove that the resulting exploration in both negotiation protocols is identical to the optimal exploration if carried out stand-alone. For the non-interleaved protocol this is straightforward: the proposer merely needs to guarantee the responder its discounted utility from continuing the negotiation by rejecting the current offer, i.e., $\delta U_{responder}(t+1,v,\mathcal{B})$. Since the sum of utilities is $EV(t,v,\mathcal{B})$ (for the same considerations discussed for the interleaved protocol), the proposer's expected utility is $EV(t,v,\mathcal{B})-\delta U_{responder}(t+1,v,\mathcal{B})$. This latter term is maximized when $EV(t,v,\mathcal{B})$ is maximized, i.e., when offering to follow the optimal exploration strategy. For the interleaved negotiation protocol the proof is inductive: when getting to the last negotiation step the responder's expected utility onwards is zero, hence the proposer will offer exploring the opportunity which yields the maximum net benefit when taking into consideration value improvement and exploration cost. Now assume that the proposer at any negotiation step $t'>t$ offers the exploration of the next opportunity to explored according to the optimal stand-alone exploration strategy. We need to show that the proposer at the negotiation step t follows the same rule. Since the proposer's expected utility is $EV(t,v,\mathcal{B})-\delta U_{responder}(t+1,v,\mathcal{B})$, the proposer will attempt to maximize $EV(t,v,\mathcal{B})$. Choosing to offer the exploration of the next opportunity according to the optimal stand-alone exploration strategy will guarantee, according to the induction assumption that the remaining exploration will also follow the optimal exploration strategy and hence that the joint utility from the resulting exploration, $EV(t,v,\mathcal{B})$, is maximized. Therefore the proposer will always follow the optimal exploration strategy.*

Since the resulting exploration in both negotiation protocols follow the optimal exploration sequence, their expected joint utility $EV(t,v,\mathcal{B})$ is identical. This enables an inductive proof for the main theorem. In the final negotiation step T, the expected utility of the responder from further exploration is zero in both protocols, and the expected joint utility from resuming the process, $EV(T,v,\mathcal{B})$, is equal. Therefore, the expected utility of the proposer in both negotiation protocols is identical. Now assume the expected utilities of the agents when reaching any state $(t'>t, v'\geq v, \mathcal{B}'\subset\mathcal{B})$ are identical in both negotiation protocols. We need to prove that the expected utilities of the agents when in state $(t,v,\mathcal{B})$ are also identical. The expected utility of the proposer in both protocols is $EV(t,v,\mathcal{B})-\delta U_{responder}(t+1,v,\mathcal{B})$ and for the responder $\delta U_{responder}(t+1,v,\mathcal{B})$. Now since the values of $EV(t,v,\mathcal{B})$ and $U_{responder}(t+1,v,\mathcal{B})$ are identical in both protocols (the first due to following the same exploration strategy as proved above and the second due to the induction assumption), so is the utility $U_P(t,v,\mathcal{B})$ of each of the agents. $\square$

Using Theorem 1 the agents can replace the interleaved negotiation with a simpler non-interleaved one that yields the same outcomes for both agents. Computing the outcomes in the new protocol requires computing the optimal exploration sequence for T stand-alone exploration problems of the form $(t,0,\mathcal{B})$. In many settings, as discussed at the beginning of this section, the optimal solution for the exploration problem is immediate and thus the complexity of solving the negotiation problem becomes linear in the number of allowed negotiation steps. Furthermore, Theorem 1 can also be used as an efficient means for computing the payments that need to be made by Agt_1 even if the interleaved protocol is preferred to be used (e.g., due to its many advantages as discussed above): on each negotiation step, the proposer will offer exploring the next opportunity according to the optimal exploration policy. The payment $M(t,v,\mathcal{B})$ will be determined as the difference: $U_{responder}(t,v,\mathcal{B})-\delta U_{responder}(t+1,v,\mathcal{B})$, where $U_{responder}(t,v,\mathcal{B})$ and $U_{responder}(t+1,v,\mathcal{B})$ are calculated as the solution to the two appropriate instances of the non-interleaved negotiation problem.

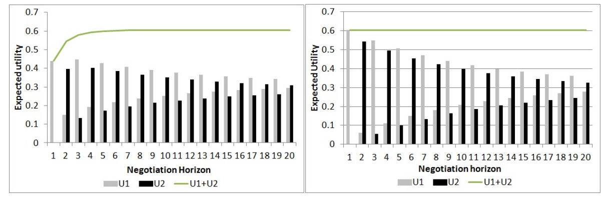
Controlling The Negotiation Scheme

In many scenarios, the agents are not symmetric in their power over the negotiations. For example, the oil drilling company may choose the time period for which it leases the franchise from the government, thus limiting the negotiation horizon. In a classic exploration setting, it will always prefer a longer franchise over a short one (assuming that it does not cost more). In this section we will show that when negotiations are involved, sometimes the company would prefer the shorter franchise, even if it costs the same. In particular, in such cases the resulting exploration process may be suboptimal and consequently the social welfare may decrease. We demonstrate some of these implications using a synthetic simplistic environment where the set of opportunities $B = \{b_1, \ldots, b_n\}$ are homogeneous, i.e., share the same exploration cost and distribution of values (uniform between zero to one, i.e., $f_i(y) = 1$), with a zero fall back ($v = 0$). We also prove some bounds on the effect of some forms of extra control, for any set of opportunities.

We begin by demonstrating the effect of the negotiation horizon over the expected utility of the two agents in the interleaved protocol. This is illustrated in Figure 1(a), along with the total joint utility (which is equal to the social welfare) as a function of the negotiation horizon (the horizontal axis). The number of initially available opportunities in this example is 20, the cost is $c = 0.01$, and discounting factor is $\delta = 0.9$. From the figure we observe that given the option to set the negotiation horizon, Agt_1 would have chosen $T = 3$ whereas Agt_2 would have chosen $T = 4$. This contrasts the results known for legacy protocols of negotiation over a single issue with complete information (Rubinstein 1982). The mapping to these protocols is based on allowing the agents to make offers that outline the negotiation as a whole and carry out the exploration as a completely separate process (without any restrictions on the number of explorations made) once a proposal is accepted. The agents thus will be simply negotiating over the division of the expected utility from an optimal exploration process, which does not change over time (except for discounting). According to the solution to these latter protocols Agt_1 will always prefer a single negotiation step and Agt_2 will always prefer two negotiation steps. This is demonstrated in Figure 1(b) for the same environment.

In addition to the above, several other differences between the two protocols exist. First, in the non-interleaved protocol (the "'single issue"'-like negotiation) the expected joint utility that the agents split between them does not change as a function of the decision horizon, and the the agents split it more equally when the number of negotiation periods increases (Sandholm and Vulkan 1999). This is demonstrated in Figure 1(b). In contrast, when the negotiation and exploration are interleaved, the joint expected utility increases as the negotiation horizon increases, because the number of opportunities that can be explored increases. For example, the use of $T = 3$ in the interleaved model results in an overall sub-optimal social welfare (joint expected utilities) for the agents as demonstrated in Figure 1(a).

The explanation for this dissimilarity is that in the interleaved protocol the deadline is not the only factor which determines the agents' utilities, but the discounting factor and cost also have an effect. In the interleaved protocol the number of opportunities is limited by the negotiation horizon, so the agents need to find the balance between increasing the negotiation horizon (potentially enabling the exploration of more opportunities) and the effect of being the last pro-



(a) Interleaved protocol (b) Non-interleaved protocol

Figure 1: Agents' utilities and joint utility.

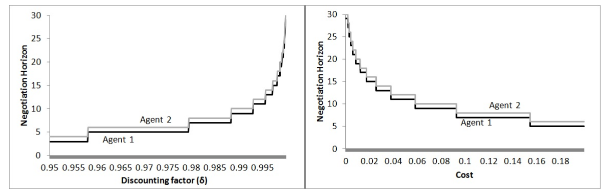


Figure 2: The effect of exploration parameters over the preferred negotiation horizon.

poser (which also depends on the discounting factor). This is further illustrated in Figure 2 which depicts the number of negotiation steps which maximizes the utility of each agent (out of 30 available negotiation steps), assuming Agt_1 is the first to propose, as a function of δ and c, respectively. In Figure 2a the cost used is $c = 0.01$ and in Figure 2b the discounting factor used is $\delta = 0.99$. As the discounting factor increases, its effect over the negotiation decreases, and the agents prefer longer negotiation horizons so that the expected joint utility divided between them increases. A similar effect occurs with costs — as the cost of exploration increases, the benefit in further exploration decreases and thus the agents prefer to use a closer decision horizon.

Another negotiation parameter an agent can exploit to increase its expected utility is the intensity of the exploration. Our protocol originally assumes that only one exploration can be carried out at a time. However, taking into account the discounting of gains and the fact that the number of steps is limited, higher social welfare (and consequently individual utilities) can be often obtained using parallel exploration (Benhabib and Bull 1983; Gal, Landsberger, and Levykson 1981).In this case the agents negotiate over the payment Agt_1 should pay Agt_2 for exploring q opportunities at a time (i.e., in parallel) with an exploration cost $q * c$.

The analysis of the parallel negotiation protocol uses backward induction similar to the one used with the interleaved negotiation protocol. The optimal exploration strategy for the corresponding stand-alone problem is once again threshold-based and can be extracted with a polynomial complexity (Morgan and Manning 1985; Morgan 1983). In this case, once again, if one of the agents is given control over the number of opportunities that can be explored over each negotiation step, it will choose that number to increase its own expected utility, which may result in a decrease in the joint utility. Figure 3 depicts the number of explorations to be executed in parallel that maximize the expected utility of Agt_1 out of the negotiation in comparison to the number of parallel explorations that maximize the expected joint utilities, as a function of δ (within the interval $0.93 - 1$).

Figure 3: Number of explorations to be executed in parallel.

The exploration cost in this example is $c = 0.01$ and the negotiation horizon is 30. In this example, only pure divisors of the negotiation horizon are considered as legitimate q values. From the figure we observe a reverse correlation between the social welfare and Agt_1's expected individual utility. Note that Agt_1 will choose q that will cause it to be the last proposer, where the identity of the last proposer is immaterial to maximizing the joint utility.

The example above suggests that if one of the agents gets control over one of the negotiation parameters, then it potentially can set it in a way that its own expected utility out of the negotiation increases at the expense of the overall social welfare. We now explore the ratio between an agent's utility with and without such control, and the ratio between the social welfare in protocol with the control and the social welfare in the protocol without the control. This lets us bound the effect of the extra control, for every set of opportunities.

Suppose that there are m identical opportunities, and that Agt_2 is capable of exploring q opportunities in parallel (by paying the cost of all of them). Let $SW(q)$ denote the social welfare when q opportunities are explored at each step (until the agents decide to stop exploring). Let q_{max} be the number of opportunities which maximizes $SW(q)$. If the agents would cooperate, they would choose to explore q_{max} opportunities at each step. Consider the setting in which Agt_1 first chooses a number q, and then at every step of the negotiations, one of the agents offers a price for exploiting q opportunities at a time. Note that Agt_1 will choose q to maximize its utility from the entire procedure, taking into account both the negotiations and the exploration. Let $U_1(q)$ denote Agt_1 utility if q opportunities are explored at each step, and let q_1 be the value which maximizes Agt_1 utility.

We now show that Agt_1 can greatly improve its utility using its control over the negotiation

Proposition 1. *The ratio $U_1(q_1)/U_1(q_{max})$ is unbounded.*

Proof. Consider a setting with two identical opportunities, discounting $\delta = 1 - \epsilon^2$ and two negotiation periods. Each opportunity gives value 1 and has cost $\frac{1}{2} - \epsilon$ for some small constant ϵ. We have that $q_{max} = 1$ because after exploring one opportunity there is no marginal benefit from the second opportunity and the exploration will be terminated. When $q = 2$ both opportunities will be explored but the highest value will remain 1. In the last negotiation step Agt_2 will gain the marginal benefit from the exploration $U_2 = 1 - (0.5 - \epsilon) = 0.5 + \epsilon$. Because the opportunities are identical then in the previous negotiation step Agt_1 will offer $M = \delta * U_2 = (1 - \epsilon^2)(0.5 + \epsilon)$ for exploring the opportunity and $U_1(1) \leq \epsilon^2$. However, one can see that $q_1 = 2$, and indeed $U_1(2) = 1 - 2(0.5 - \epsilon) = 2\epsilon$. The ratio between $U_1(2)/U_1(1)$ is unbounded as $epsilon$ approaches zero. $\square$

Noting that in the previous example $SW(1) = \frac{1}{2} + \epsilon$ but $SW(2) = 2\epsilon$ proves the following proposition:

Proposition 2. *The ratio $SW(q_1)/SW(q_{max})$ is unbounded or equivalently the Price of Anarchy in this setting is unbounded.*

Note also that in this example the utility of Agt_2 drops from $\frac{1}{2} - \epsilon$ to ϵ.

We now consider a setting with heterogeneous opportunities in which Agt_1 chooses exactly what opportunities will be negotiated over at each step. Surprisingly, this setting behaves differently than the previous one, in which Agt_1 was only able to choose once how many opportunities will be negotiated in each step (but it had to be the same number every step). Let $SW_{OPT}(b_1, \ldots, b_n, \delta)$ denote the maximal social welfare of the exploration problem, when both agents cooperate during the negotiation. Let $SW_{1choose}(b_1, \ldots, b_n, \delta)$ denote the social welfare when Agt_1 chooses which opportunities will be negotiated at each step, to maximize its utility. The following proposition shows that the Price of Anarchy of this setting is bounded:

Proposition 3. *The welfare loss in this setting is bounded, $SW_{1choose}(b_1, \ldots, b_n, \delta) \geq SW_{OPT}(b_1, \ldots, b_n, \delta^2)$*

Proof. Consider a setting with opportunities $\{b_1, \ldots, b_n\}$ and discounting factor δ. We show a strategy which guarantees Agt_1 a utility of $SW_{OPT}(b_1, \ldots, b_n, \delta^2)$. This means that the maximal utility for Agt_1 will be at least $SW_{OPT}(b_1, \ldots, b_n, \delta^2)$. As the social welfare of the protocol is at least the utility of Agt_1, this proves the proposition. Let S be an optimal strategy for the exploration of $\{b_1, \ldots, b_n\}$ if the discounting is δ^2. At the odd stages, when Agt_1 makes the offers, the negotiation will be on the set of opportunities which S would utilize. In the even stages, when Agt_2 makes the offers, the negotiations will be over an empty set of opportunities. It is easy to see that if Agt_1 uses this strategy, then Agt_2 will get no utility at all, and Agt_1 will get a utility of $SW_{OPT}(b_1, \ldots, b_n, \delta^2)$, as required. $\square$

In the interleaved repetitive negotiation protocol there is a combination of negotiation and exploration, and in this section we demonstrate that one influence the other, and by changing a single parameter in one of the settings the agents' strategies and the social welfare optimality will be affected.

Conclusions and Future Work

As discussed throughout the paper, a negotiation scheme which interleaves negotiation and exploration has several important inherent advantages and is likely to be preferred in some real-life settings. Yet, the new protocol is also associated with a substantial increased computational complexity. Theorem 1 enables us both to determine the exploration that will be offered and compute the payments on each step in the new protocol, which allows us to analyze some of its key features. In particular, the paper demonstrates that if one of the agents gets partial control over the protocol parameters, different preferences may be revealed in comparison to preferences known in non-interleaved negotiation. The implication of this latter result concerns mainly social welfare

— an agent may choose to deviate to a sub-optimal (social-welfare-wise) negotiation if it increases its own expected utility. The degree of control the agent obtains has a substantial effect on the degradation in welfare. For example, if the agent is allowed to choose once how many opportunities will be explored at every step, there can be a sharp drop in social welfare, where if the agent can change this number at every step, the drop is bounded. As in real life situations it is common for one agent to have some limited control over the parameters (consider the oil drilling example, the NSF which chooses in which topics to give grants etc.), continuing this exploration is an important open problem.

Finally, we note that in some domains negotiation cycles can be substantially faster to carry out in comparison to exploration. In such case protocols where several negotiation steps take place, followed by exploration, can be considered. The analysis methodology still holds in this case. Also, we emphasize that the proof given in Theorem 1 can be generalized to other settings (not necessarily exploration-based, e.g., a house owner and a home improvement contractor) where multiple tasks can be negotiated and the agents prefer to negotiate and execute the tasks agreed upon prior to continuing the negotiation. The analysis of such domains is, of course, beyond the scope of the current paper and is thus left to future research.

Acknowledgments

This work was supported in part by ISF/BSF grants 1401/09 and 2008-404.

References

Bac, M., and Raff, H. 1996. Issue-by-issue negotiations: the role of information and time preference. *Games and Economic Behavior* 13(1):125–134.

Benhabib, J., and Bull, C. 1983. Job search: The choice of intensity. *J. of Political Economy* 91(5):747–764.

Buttner, R. 2006. A classification structure for automated negotiations. WI-IAT '06, 523–530.

Chevaleyre, Y.; Endriss, U.; and Maudet, N. 2010. Simple negotiation schemes for agents with simple preferences: Sufficiency, necessity and maximality. *Autonomous Agents and Multi-Agent Systems* 20(2):234–259.

Chhabra, M.; Das, S.; and Sarne, D. 2011. Expert-mediated search. AAMAS '11, 415–422.

Dunne, P.; Wooldridge, M.; and Laurence, M. 2005. The complexity of contract negotiation. *Artificial Intelligence* 164(1-2):23–46.

Faratin, P.; Sierra, C.; and Jennings, N. R. 2002. Using similarity criteria to make issue trade-offs in automated negotiations. *Artificial Intelligence* 142:205–237.

Fatima, S.; Wooldridge, M.; and Jennings, N. 2004. An agenda-based framework for multi-issue negotiation. *Artificial Intelligence* 152(1):1–45.

Fatima, S.; Wooldridge, M.; and Jennings, N. 2006. Multi-issue negotiation with deadlines. *JAIR* 27(1):381–417.

Fershtman, C. 1990. The importance of the agenda in bargaining. *Games and Economic Behavior* 2(3):224–238.

Gal, Y.; Kraus, S.; Gelfand, M.; Khashan, H.; and Salmon, E. 2011. An adaptive agent for negotiating with people in different cultures. *ACM Transactions on Intelligent Systems and Technology (TIST)* 3(1):8.

Gal, S.; Landsberger, M.; and Levykson, B. 1981. A compound strategy for search in the labor market. *International Economic Review* 22(3):597–608.

Jennings, N.; Faratin, P.; Lomuscio, A.; Parsons, S.; Wooldridge, M.; and Sierra, C. 2001. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation* 10(2):199–215.

Kraus, S. 1997. Negotiation and cooperation in multi-agent environments. *Artificial Intelligence* 94:79–97.

Larson, K., and Sandholm, T. 2002. An alternating offers bargaining model for computationally limited agents. AAMAS '02, 135–142.

Lin, R., and Chou, S. 2004. Mediating a bilateral multi-issue negotiation. *Electronic Commerce Research and Applications* 3(2):126–138.

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.

Lomuscio, A.; Wooldridge, M.; and Jennings, N. 2001. A classification scheme for negotiation in electronic commerce. *Agent Mediated Electronic Commerce* 19–33.

Ma, C.-T. A., and Manove, M. 1993. Bargaining with deadlines and imperfect player control. *Econometrica* 61(6):1313–39.

McMillan, J., and Rothschild, M. 1994. Search. In Aumann, R., and Hart, S., eds., *Handbook of Game Theory with Economic Applications*. 905–927.

Morgan, P., and Manning, R. 1985. Optimal search. *Econometrica: Journal of the Econometric Society* 923–944.

Morgan, P. 1983. Search and optimal sample sizes. *The Review of Economic Studies* 50(4):659.

Rosenschein, J., and Zlotkin, G. 1994. *Rules of encounter: designing conventions for automated negotiation among computers*. the MIT Press.

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

Rubinstein, A. 1985. A bargaining model with incomplete information about time preferences. *Econometrica* 53(5):1151–72.

Sandholm, T., and Vulkan, N. 1999. Bargaining with deadlines. In *In AAAI-99*, 44–51.

Sarne, D.; Manisterski, E.; and Kraus, S. 2010. Multi-goal economic search using dynamic search structures. *Autonomous Agents and Multi-Agent Systems* 21(2):204–236.

Stahl, I. 1972. *Bargaining theory*. (Economics Research Institute, Stockholm School of Economics, Stockholm).

Stigler, G. 1961. The economics of information. *Journal of Political Economy* 69(3):213–225.

Weitzman, M. L. 1979. Optimal search for the best alternative. *Econometrica* 47(3):641–54.