2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Workshops

The Benefits of Opponent Models in Negotiation

Koen Hindriks, Catholijn M. Jonker and Dmytro Tykhonov
Man-Machine Interaction, Faculty of EEMCS, Delft University of Technology,
Mekelweg 4, 2628CD Delft, The Netherlands
{k.v.hindriks, c.m.jonker, d.tykhonov}@tudelft.nl

***Abstract*— Information about the opponent is essential to improve automated negotiation strategies for bilateral multi-issue negotiation. In this paper we propose a negotiation strategy that exploits a technique to learn a model of opponent preferences in a single negotiation session. An opponent model may be used to achieve at least two important goals in negotiation. First, it can be used to recognize, avoid and respond appropriately to exploitation, which differentiates the strategy proposed from commonly used concession-based strategies. Second, it can be used to increase the efficiency of a negotiated agreement by searching for Pareto-optimal bids. A negotiation strategy should be efficient, transparent, maximize the chance of an agreement and should avoid exploitation. We argue that the proposed strategy satisfies these criteria and analyze its performance experimentally.**

Multi-issue negotiation; opponent modelling; Bayesian learning; negotiation strategy; Tit-for-Tat

I. INTRODUCTION

In bilateral negotiation, two parties aim at reaching a joint agreement, by exchanging various offers using e.g. an alternating offers protocol. Two basic, constitutive facts about negotiation define the basic dilemma each negotiator has to face: (1) each party aims to satisfy its own interests as best as possible, but (2) in order to reach an agreement one has to take ones opponent's preferences into account as well.

In the literature on automated negotiation, typically, concession-based strategies have been proposed. An agent that uses a concession-based strategy selects as the next offer it will make an offer that has a decreased utility compared with the last offer made. The utility that is being decreased is the utility from the agent's own perspective without any guarantee that such a decrease will also increase the utility from the other party's perspective. A well-known example of such a strategy is the time-dependent strategy which decreases utility simply as a function of time [6]. Although motivated by fact (2) above, such strategies do not explicitly take the opponent's preferences into account, and, as a result, will most likely be inefficient in complex negotiation domains. Moreover, time-dependent strategies can be exploited by the other negotiating party and as such do not adequately take fact (1) above into account.

The solution to these problems is to explicitly take the preferences of an opponent into account. The benefits of doing so are that it enables a search through the outcome space for outcomes that are mutually beneficial and that it allows classifying and recognizing the type of move an opponent has made. In order to do so, two key questions need to be addressed: How can an agent obtain information about opponent preferences? And: How can an agent exploit information about opponent preferences effectively?

In this paper we consider single session negotiations, i.e., negotiators cannot learn from repeated sessions with the same opponent. As negotiators typically are not willing to reveal their preferences to avoid exploitation, information about opponent preferences needs to be obtained from the behaviour of that opponent. The first question is addressed by means of opponent modelling techniques, several of which have been proposed, see e.g. [2, 13]. We use a technique based on Bayesian learning here that is able to effectively learn opponent preferences during a single negotiation session [4].

This paper shows how opponent preferences can be strategically exploited in negotiation. It is organized as follows. Section 2 discusses related work. In Section 3 a design of a negotiation strategy that explicitly uses opponent preferences is introduced. The theme of Section 4 is the algorithm of the proposed negotiation strategy. Its effectiveness is validated in Section 5 by way of experimental results. Section 6 concludes the paper.

II. RELATED WORK

In this Section we first discuss related work on negotiation strategies, and then we briefly discuss related work on learning and introduce the technique we used in our experiments.

The literature on negotiation strategies is extensive and we only discuss some examples to illustrate the variety of ideas that have been proposed to design such strategies. In [6] a range of decision functions that may be used to define a negotiation strategy are discussed, focussing on different aspects that may be relevant in a negotiation such as time and the behaviour of an opponent. As no single tactic or decision function seems to be "right" in arbitrary negotiation settings, an approach proposed in [10] aims at combining different types of such negotiation tactics from [6] in a single strategy. An evolutionary algorithm is used to compute a next offer that adjusts the weights associated with each of the individual tactics. This approach is not suitable for the one-session closed negotiation situation we are focussing on. To begin with it requires a substantial number of negotiations to learn appropriate weights associated with the tactics. Moreover, the preference profiles of both parties must be made public in order to calculate the fitness during the

978-0-7695-3801-3/09 $26.00 © 2009 IEEE
DOI 10.1109/WI-IAT.2009.192

learning phase. As a result, the weights learned to combine the strategies only yield efficient negotiations in negotiation setups that are one-session and closed.

The negotiation strategy of [7] can be used in one-session closed negotiation setting. It postpones concessions by using domain knowledge to offer bids that increase the utility (and thus acceptability) for the opponent without decreasing the utility associated with the offer for the agent making the offer. Such offers increase the chances of a proposed trade-off that is good for both parties. Nevertheless, this strategy concedes even if the opponent does not.

The concession-based negotiation strategy of [8] determines the size of its next concession mainly on the basis of the utility gap between the last bids of the agent and the opponent. The next bid configuration, however, is based purely on the agent's own preference profile and thus reaches no win-win outcomes. Its time-dependent nature means that it can be exploited by the opponent.

To tackle the problem of exploitation, in [6], a number of variants of Tit-for-Tat tactics are discussed that belong to the family of so-called behaviour-dependent tactics. These tactics, however, do not use an opponent model and only vary utility of the agent's own perspective consistently. These tactics thus are blind to the preferences of an opponent. The behaviour generated by such a tactic therefore is not transparent and may be hard to understand from the opponent's point of view. With Axelrod [1984], we consider transparency an important feature of any strategy, which has motivated the design of the strategy introduced here. Transparency may be achieved by using available knowledge about the preference profile of the opponent, as explained in Section 3.

Various approaches to learning in a negotiation context have been based on forms of Bayesian learning, e.g. [4, 13]. For the negotiation situations we are focusing on, we need a technique that is able to learn the opponent profile during one session, such as the Bayesian learning technique introduced in [4], which we chose as a building block in this paper.

III. NEGOTIATION STRATEGY DESIGN

The preferences of an opponent can be used in at least two ways. First, it can be used to propose efficient Pareto-optimal offers. Finding such offers requires that the Pareto frontier can be approximated which is only feasible if a reasonable model of the opponent's preferences is available. Second, it can be used to recognize and avoid exploitation. The strategy we propose is inspired by a classification of negotiation moves as described in [3] and the Tit-for-Tat tactic, discussed in [1] and – in a negotiation context – in [6]. As learning techniques will not provide perfect models of an opponent's preferences a strategy should be robust with respect to such imperfections. We return to this last point in the Section 5. The design of the negotiation strategy proposed in this paper is based on a number of observations and criteria that we want the strategy to satisfy. The main criteria are that the strategy should be *efficient, transparent, maximize the chance of an agreement and should avoid exploitation.*

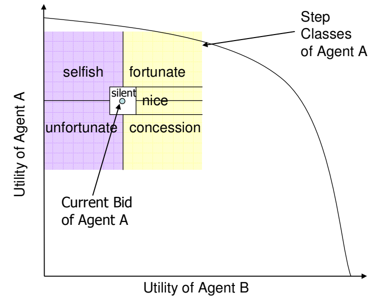

Figure 1. Classification of negotiation moves

The first observation relevant to the design of our strategy is that the availability of information about the preferences of an opponent enables an agent to classify the moves its opponent makes. Here, we use a classification of moves proposed in [3] and illustrated in Figure 1. The move classification is presented from the perspective of agent A.

Given that agent A's last offer is marked by the arrow "Current Bid of Agent A", the agent has a number of choices for making a next negotiation move. A *silent move* does not change utility of either party significantly. A *concession move* decreases own utility but increases the utility of the opponent. A *fortunate move* increases utility for both parties whereas an *unfortunate move* does the opposite. Note that a fortunate move can only be made if the current bid is not already on the Pareto frontier. A *selfish move* increases own utility but decreases the opponent's utility. Finally, a *nice move* increases the opponent's utility but does not change the agent's own utility.

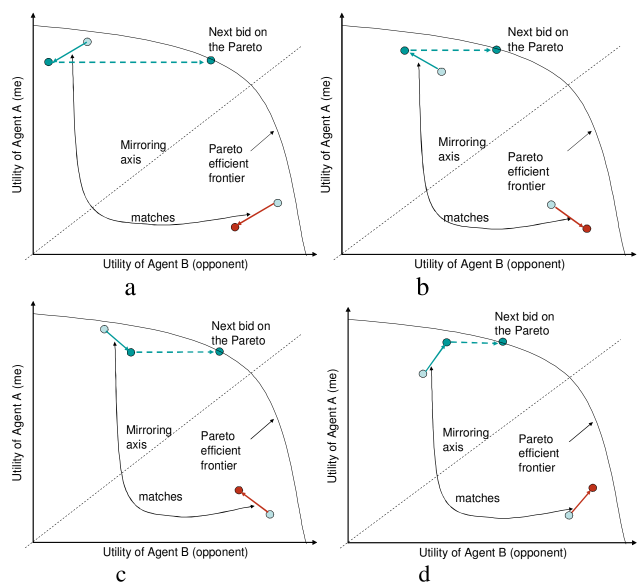

Figure 2. Example responses for an (a) unfortunate move, (b) selfish move, (c) concession move, and (d) fortunate move.

Based on this classification a simple suggestion would be to "mirror" each move of an opponent by making a similar move, which would implement a Tit-for-Tat-like tactic. The basic idea of a Tit-for-Tat strategy in a multi-issue negotiation context would be to respond to an opponent move with a symmetrical one. That is, "match" the move as depicted in Figure 2 by mirroring it in the diagonal axis.

First note that each type of move would indeed result in a response move in the same class. In particular, responding to a *concession move* of the opponent with a *concession move* itself arguably is one of the most reasonable responses one can make. All rational negotiation strategies will attempt to make concession moves at some point during a negotiation. Moreover, the "mirroring" strategy would avoid exploitation as a selfish move of the opponent would result in a selfish response move. Such a response would be a *signal* to the opponent: "I am prepared to make a concession towards you only if I get something in return. If you pull back I'll do the same".

A mirroring strategy would, however, be too simplistic for several reasons. A mirroring strategy is not rational in the case of an unfortunate move, as there is no reason to decrease the agent's own utility without increasing the chance of acceptance of the proposed bid by the opponent. Furthermore, observe (compare Fig. 2) that unfortunate moves move away from the Pareto-optimal frontier, and thus would not satisfy our efficiency criteria.

In order to remove these deficiencies, we propose to first mirror the move of the opponent and thereafter make an additional move towards the Pareto frontier, i.e. a move towards the approximated Pareto frontier that is computed using the learned opponent model and the agent's own preference profile. There are multiple ways to do this and the choice is not straightforward. What is clear is that the move towards the Pareto frontier should not further *decrease* the agent's own utility as this would invite exploitation tactics. Furthermore, it also does not seem rational to further *decrease* the opponent's utility as this would result in *selfish moves* to arbitrary moves of the opponent.

The final observation that motivated our choice is that *increasing* the agent's own utility by moving towards the Pareto frontier actually *minimizes* the chance of reaching an agreement when this strategy would be used by both parties, which would violate one of our design criteria for a negotiation strategy. To explain this, consider two agents that would mirror an opponent's move and then, seen from the perspective of Agent A in Figure 2, would move straight up towards the Pareto frontier (Agent B would move right) which would *only* increase own utility. The other agent in this case would consider such a move a selfish move and respond similarly, thereby minimizing the chance of reaching an agreement. Of course, this line of reasoning depends on the quality of the opponent model but presents a real problem. To resolve it, the strategy we propose only increases the opponent's utility when moving towards the Pareto frontier in order to maximize the chance of an agreement. The resulting strategy consists of two steps: first *mirror the move* of the opponent and then add a *nice move* to propose an efficient offer (i.e., search for a bid on the approximated Pareto frontier that is on the same iso-curve as the bid obtained by mirroring, see Figure 2). This strategy we call the *Mirroring Strategy* (MS).To gain a better understanding of MS, it is instructive to discuss some of the response moves MS generates. Figure 2 shows examples of responses to an unfortunate, selfish concession and fortunate move. The response to an unfortunate move is to mirror this move and add a nice move, which results in a concession move (see Figure 2a). This is a reasonable reply as such a move may be interpreted as an attempt (that failed) to make a concession move by the opponent (due to the lack of information about the preferences of its opponent). Such a move which is the result of misinformation should not be punished, we believe, but an attempt instead should be made to maintain progress towards an agreement.

The response to a *selfish move* either results in a *fortunate move* or in a *selfish move*. Figure 2b shows the case resulting in a fortunate move. It should be noted that a fortunate move is only possible if the previous move the agent made was inefficient. This means that in that case the opponent model must have misrepresented the actual preferences of the opponent. In such a case, where our previous move was based on misinformation, we believe it is reasonable to not punish the opponent with a selfish move and give the opponent the benefit of the doubt in such a case. If, however, the previous move would have been efficient, a selfish move most likely would be replied to with a selfish move (since there would be no room to make a nice move towards the Pareto frontier), and it is reasonable to send a clear signal to the opponent that such moves are unacceptable.

Finally, both a *concession move* as well as *an unfortunate move* of the opponent would be replied to with the same type of move (see Figure 2c and 2d). Moreover, if there is room for a nice move towards the Pareto frontier, in both cases the step would be bigger than that made by the opponent, increasing the utility of the opponent even more and thereby again increasing the chance of acceptance as early on in a negotiation as possible.

As discussed, a negotiation strategy should be *efficient, transparent, maximize the chance of an agreement* and should *avoid exploitation*. It is clear that MS aims to be as efficient as possible, which only depends on the quality of the learning technique for modelling opponent preferences. Performance of the learning algorithm used in MS was studied in [5]. The study concludes that the learning algorithm can learn the most important aspects of the opponent preferences in a range of negotiation settings. MS does not aim at exploiting the weaknesses of an opponent strategy. Instead it aims for restoring efficiency whenever an opponent strategy is not able to do so and aims at a fair outcome (see also Section 5 below). MS is transparent as it is proposes a simple response strategy by mirroring an opponent's move and then adding a nice step. The signals thus send by negotiation moves are easy to interpret by an opponent. In particular, MS only punishes an opponent in reply to a selfish move and only does so when the model of opponent preferences matches the actual preferences of that opponent. As a result, MS not only avoids exploitation but also is a nice strategy. MS is nice even when an opponent makes unfortunate moves which are interpreted as "mistakes" on the opponent's part. The strategy moreover maximizes the chance of an agreement as early as possible, which is achieved by the move towards the Pareto frontier that always maximizes the utility of the opponent relative to a particular utility for the agent itself.

IV. MATHCING STRATEGY ALGORITHM

Here we present the MS strategy in 6 algorithmic steps, including the steps needed for learning an opponent model. The algorithm is presented in Figure 3. As is usual, MS starts by proposing a bid that has maximal utility with respect to the agent's own preference profile (step 1). In step 2 a simple but reasonable acceptance strategy is used which is not particular to MS. A bid from an opponent is accepted when the utility of that bid is higher than that of the agent's own last bid or the utility of the bid it would propose next. Otherwise, the agent will propose a counter-offer. In step 3, the bid received from the opponent is used to update the opponent model with the function $\tilde{U}(\omega)$ of [4]. Steps 4, 5 and 6 define MS. Step 5 mirrors an opponent's move, after which step 6 determines a nice move towards the Pareto frontier (given the opponent model computed in step 3).

1	Start with a bid of maximal utility w.r.t. the agent's own preference profile: $$\omega_0 = \arg\max_{\omega \in \Omega} U(\omega)$$
2	Receive an opponent's bid ω^o_t. Accept the opponent's bid if its utility is at least high as the utility of own previous bid: $$action = \begin{cases} accept, & U(\omega^o_t) \geq U(\omega_{t-1}) \\ goto\ step\ 3, & otherwise \end{cases}$$
3	Update the opponent preference model given : $$\omega^o_t : \tilde{U}_{t-1}(\omega) \xrightarrow{update} \tilde{U}_t(\omega)$$
4	Classify the opponent's move ω'_{t-1} to ω'_t: *concession :* $\Delta U(\omega'_t) > 0$ *and* $\Delta\tilde{U}(\omega'_t) < 0$ *unfortunate:* $\Delta U(\omega'_t) < 0$ *and* $\Delta\tilde{U}(\omega'_t) < 0$ *selfish:* $\Delta U(\omega'_t) < 0$ *and* $\Delta\tilde{U}(\omega'_t) > 0$ *fortunate:* $\Delta U(\omega'_t) > 0$ *and* $\Delta\tilde{U}(\omega'_t) > 0$ where: $$\Delta U(\omega'_t) = \Delta U(\omega'_t) - \Delta U(\omega'_{t-1})$$ $$\Delta\tilde{U}(\omega'_t) = \Delta\tilde{U}(\omega'_t) - \Delta\tilde{U}(\omega'_{t-1})$$
5	Calculate coordinates of the matching bid ω_t', given the classification in the previous step: *fortunate, unfortunate:* $\Delta U(\omega^o_t) = \Delta U(\omega'_t)$; $\Delta\tilde{U}(\omega^o_t) = \Delta\tilde{U}(\omega'_t)$ *selfish, concession:* $\Delta U(\omega^o_t) = -\Delta U(\omega'_t)$; $\Delta\tilde{U}(\omega^o_t) = -\Delta\tilde{U}(\omega'_t)$ Then the coordinates of ω_t' are: $$U(\omega'_t) = U(\omega'_{t-1}) + \Delta U(\omega'_t)$$ $$\tilde{U}(\omega'_t) = \tilde{U}(\omega'_{t-1}) + \Delta\tilde{U}(\omega'_t)$$ If $U(\omega'_t)$<*reservation value* then stop negotiation
6	Find a bid ω_t that corresponds to the ω'_t and belongs to the approximated Pareto frontier and belongs to the same iso-curve on the $U(\omega'_t)$: $$\omega_t = \arg\max_{\omega \in \Omega, U(\omega) = U(\omega'_t)} \tilde{U}(\omega)$$
7	Send bid ω_t to the opponent.
8	Go to step 2.

Figure 3. Matching Strategy algorithm

Note, that in the beginning of a negotiation the model of the opponent preferences is inaccurate because it has not been updated yet. However, this is not crucial at this stage of the negotiation given that both agents would start by offering a bid with the highest utility for themselves that is the most rational choice. The initial offer of the opponent is used to make the first update of the opponent model. The second move of the MS strategy can only be a concession or an unfortunate move because the initial offer cannot be improved for either player and, therefore, the other classes of moves are not possible. In this case the size of the first concession in the opponent's utility would be determined by the efficiency of the learning technique. Given the fact that MS always tries to maximize the opponent's utility it will try to make a concession move, thus signalling to the opponent its readiness to proceed in the same manner.

The MS strategy is developed to avoid exploitation by the state-of-the-art rational strategies and tries to match the opponent's moves in a transparent way as it is defined by the design criteria. The MS strategy is experimentally tested in a tournament setting against such strategies (see Section 5). To prevent exploitation by irrational strategies we use a reservation value to limit the concessions made by the MS strategy. The reservation value is defined in the user's preference profile and represents a utility value below which all bids are unacceptable for the agent.

V. EXPERIMENTAL ANALYSIS

We test the efficiency of the MS strategy in an experimental setup, in which the MS strategy negotiates against automated negotiation strategies available in the literature and against human negotiators. Furthermore, this Section shows that the MS strategy constructed results in a fair agreement for both parties.

As the negotiation setup influences the negotiation performance of a strategy [3], a tournament was set up to test MS in various negotiation scenarios consisting of a domain and a preference profile for both parties. In the tournament, a range of strategies negotiated on different domains against each other. The negotiation sessions were conducted independently from each other and the results were used to evaluate the performance of MS. Each strategy negotiated 50 times against the ZI strategy (see below) and 10 times against other strategies in each scenario.

Performance of a strategy may be influenced by various features of the domain, including the level of opposition, the size of the domain, and whether preferences over issues are predictable or not (cf. [3]). Accordingly, we have selected domain and preference profiles that vary with respect to these features, summarized in Table 1. The level of opposition has been measured with respect to the importance associated with issues and with respect to the complete ranking of all possible outcomes. An issue, such as price, is called predictable when it is possible to reliably estimate the preference structure associated with the issue by an opponent. The number of issues that are unpredictable is an indicator of how difficult it is to learn an opponent model in a particular domain.

TABLE I. SUMMARY OF NEGOTIATION DOMAINS

Negotiation Domain	Opposition of preferences		Domain Size	Nr. of predictable issues
	Profile	Weights		
Car	0.64	0.60	18,750	1 (5)
Party	0.54	0.46	3,125	0 (5)
Service orient.	0.67	0.83	810,000	4 (4)
AMPOvsCity	0.66	0.42	7,128,000	3 (10)
Empl. contr.	0.70	0.60	3,125	5 (5)

The *AMPO vs City* domain [12] is the largest domain in our test. The *Party* domain developed by us is small with rather cooperative preference profiles. Humans tend to perform well on this domain. The *Service-Oriented Negotiation* (SON) domain was taken from [7]. The *Employment contract negotiation* domain was taken from [11]. The *2nd hand car* domain was taken from [8].

Besides MS four other strategies were used. The *Zero Intelligence* (ZI) strategy randomly proposes bids above its break-off point, which was set to 0.6 in the tournament. It is difficult for the ZI strategy to achieve a better agreement than its break-off point and any effective negotiation strategy should be expected to at least outperform it. We use the ZI strategy as a baseline. It also provides a good test case for any learning technique. Details about the *ABMP* strategy can be found in Section 2 and in [8], for the *Trade-off* strategy in [7] and Section 2. The *Bayesian Smart* strategy is similar to the Trade-off strategy but uses the Bayesian learning algorithm from [4] to model opponent preferences. As the same learning algorithm has been used by MS, the Bayesian Smart strategy can be used to compare performance of MS with that of the Bayesian Smart. To analyze the robustness of MS in negotiations against humans an experiment was setup in which 42 subjects first negotiated face-to-face and then negotiated against MS that used the same profile as the human opponent in the first session. The Party domain was used for the experiment. The human subjects were able to familiarize themselves with a negotiation environment used in the experiment and practice on other domains.

For every negotiation domain and preference profile, the utilities of agreements obtained by a strategy against all other strategies in the tournament were averaged. The ZI strategy was used as a baseline. Table 2 reports the percentage increase compared to the average utility of this strategy.

MS shows improved performance compared with the benchmark Bayesian Smart strategy on all domains. The

TABLE II. UTILITY INCREASE RELATIVE TO THE ZI STRATEGY

Negotiation Domain	Negotiation Strategy			
	ABMP	Trade-Off	Bayesian Smart	NMS
Car	16%	12%	13%	14%
Party domain	13%	9%	13%	14%
Service Oriented	14%	17%	25%	38%
AMPO vs City	10%	13%	14%	20%
Employment contr.	11%	40%	44%	47%

main reason is that MS is more robust since it matches the moves of its opponent and does not concede more than its opponent. The results show that on all domains MS outperforms the other strategies, except for the 2nd hand car domain where ABMP performs best. The differences on this and the Party domain are not big for all strategies.

The most significant improvement compared to ZI is achieved in the Employment contract domain. This domain is relatively small and issues are predictable. Learning an opponent model is relatively easy, and important in this domain as it contains compatible issues (i.e., both agents have similar preferences with regard to such an issue).

The results on the SON and AMPO vs City domain are comparable to that of the Employment domain. It is more difficult to reach efficient agreements in the SON domain as this domain is bigger and the variation of issue importance is much bigger. The performance of MS on the AMPO vs City domain is not as good mainly due to the decreasing performance of the learning technique in domains of high dimensionality. The improvement over the benchmark Bayesian Smart strategy is still significant in both these domains which shows that MS is a robust strategy even when the model of the opponent's preferences is not very good. Improvement is caused by the fact that MS tries to match the opponent's moves, which, at least with respect to own utility it is always able to do. Therefore, even if the quality of the learned model is low as it is e.g. in the AMPO vs City domain, MS unlike the Bayesian Smart strategy will concede only if the opponent does so too.

To analyze the robustness of MS more precisely we consider the results on the SON domain as shown in Table 3. The quality of learning is high in this domain and, therefore, is a good choice to test the robustness of MS against various strategies. Table 3 lists average utility values of the agreements reached for each party in the tournament on the SON domain. We have used the standard deviation of these utilities as a measure of the robustness of MS. The average utility value of agreements is high and the deviation of the utility of agreements is lower for MS than other strategies, which confirms that MS is more robust and more difficult to exploit than these strategies.

The technique used by the Trade-Off strategy to match the opponent's preferences strongly depends on the efficiency of the strategy used by the opponent, see [3]. E.g., the Trade-Off strategy is not able to find Pareto efficient offers in settings where it negotiates against less efficient strategies such as the ABMP and the ZI strategy. As a result, the utility of outcome reached by the Trade-Off strategy is in average higher than that of the ZI and ABMP strategy, but its deviation is relatively high (see Table 3). The learning technique used in the MS strategy, on the other hand, does not depend on the efficiency of the opponent's strategy (see [5]) and, therefore, is able to achieve better results in negotiations against the ZI and the ABMP strategy.

TABLE III. UTILITY OF AGREEMENT IN THE SON DOMAIN

Role	Utility Statistics	ZI	ABMP	Trade-Off	Bayesian Smart	NMS
Agent A	Average	0.574	0.657	0.726	0.748	0.769
	Std. dev.	0.056	0.079	0.103	0.023	0.020
Agent B	Average	0.519	0.657	0.652	0.800	0.805
	Std. dev.	0.030	0.066	0.134	0.044	0.043

Furthermore, we report on the performance of MS in an experiment with humans. The performance of MS in an experiment with human subjects also shows it is a robust strategy. Human subjects were not able to exploit NMS on the Party domain. Overall human performance was very good and close to the Pareto frontier due to simplicity the domain. Even so the humans had an advantage of training in the first face-to-face negotiation session, still MS managed to improve average utility with 5%, whereas in 30% of the experiments the increase in utility was larger than 5%.

TABLE IV. AVERAGE EUCLIDEAN DISTANCE FROM AGREEMENT TO KALAI-SMORODINSKY AND NASH SOLUTIONS

Negotiation Domain	Negotiation Strategy				
	ZI	ABMP	Trade-Off	Bayesian Smart	NMS
Distance to Kalai-Smorodinsky solution					
Car	0.20	0.13	0.12	0.12	0.11
Party domain	0.23	0.15	0.13	0.12	0.12
Service Oriented	0.25	0.23	0.16	0.14	0.11
AMPO vs City	0.20	0.15	0.13	0.13	0.12
Employment contr.	0.26	0.26	0.14	0.14	0.14
Distance to Nash Point					
Car	0.19	0.15	0.14	0.14	0.13
Party domain	0.20	0.19	0.15	0.13	0.13
Service Oriented	0.26	0.26	0.19	0.17	0.16
AMPO vs City	0.23	0.24	0.20	0.18	0.17
Employment contr.	0.26	0.26	0.14	0.14	0.14

Finally, as MS tries to match the moves of an opponent, it is reasonable to assume that MS typically tends to result in a fair outcome. This hypothesis is confirmed by Table 4, which shows that the agreements reached by MS are, on average, closer to the Nash and Kalai-Smorodinky solutions on all domains. The results also show that MS prefers the Kalai-Smorodinsky over the Nash solution.

VI. CONCLUSIONS

The Nice Mirroring Strategy introduced here shows that two important goals in closed multi-issue negotiations can be achieved when a reasonable estimate of the preferences of an opponent is available. Using a learning technique to obtain such an opponent model, it is possible to increase the efficiency of the negotiated agreement and to avoid exploitation by the other party.

The design of MS has been based on a classification of negotiation moves developed for the analysis of negotiation strategies, see [3]. MS satisfies several design criteria we believe are important for any negotiation strategy. A negotiation strategy should be efficient, transparent, maximize the chance of an agreement and should avoid exploitation. MS has been shown to be efficient and fair as it is biased towards the Kalai-Smorodinsky solution. MS is transparent as it is proposes a simple response strategy by mirroring an opponent's move and then adding a nice step, i.e. a move over the utility iso-curve towards the Pareto frontier. In fact, MS is a "nice" strategy as it will only punish an opponent in reply to a selfish move and only does so when the model of opponent preferences matches the actual preferences of that opponent. The strategy moreover maximizes the chance of an agreement as early as possible, which is achieved by the move towards the Pareto frontier, that always maximizes the utility of the opponent relative to a particular utility for the agent itself. The effectiveness of MS has been validated experimentally in a tournament setup, using domains of different characteristics and a number of different negotiation strategies. The results show that MS is able to realize significant increases in utility. In future work, we plan to investigate the exploitability of strategies and in particular MS. An important related theme for future work concerns acceptance criteria used by negotiation strategies.

REFERENCES

[1] R. Axelrod. The Evolution of Cooperation. Basic Books, Inc., Publishers, New York, USA, 1984.

[2] Coehoorn, R.M., Jennings N.R. Learning an Opponent's Preferences to Make Effective Multi-Issue Negotiation Trade-Offs, In: Proc. of 6th Int. Conference on E-Commerce, pp. 59-68, 2004.

[3] K. Hindriks, C.M. Jonker, D. Tykhonov. Analysis of Negotiation Dynamics, In: Klusch, M.; Hindriks, K.; Papazoglou, M.P.; Sterling, L. (eds.) Proc. of 11th Int. Workshop on Cooperative Information Agents, Springer-Verlag, pp. 27-35, 2007..

[4] K. Hindriks and D. Tykhonov. Opponent Modelling in Automated Multi-Issue Negotiation, In: Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS'08), Padgham, Parkes, Müller and Parsons (eds.), pp.331-338, 2008.

[5] Koen V. Hindriks and Dmytro Tykhonov, Towards a Quality Assessment Method for Learning Preference Profiles in Negotiation, In: Proc. of the 10th Int. Workshop on Agent Mediated Electronic Commerce (AMEC'08), pp. 5-17, 2008.

[6] P. Faratin, C. Sierra, and N.R. Jennings. Negotiation Decision Functions for Autonomous Agents, in: *Int. Journal of Robotics and Autonomous Systems*, 24 (3-4) 159-182, 1998.

[7] P. Faratin, C. Sierra, and N.R. Jennings. Using Similarity Criteria to Make Negotiation Trade-Offs, In: Journal of Artificial Intelligence, 142 (2), pp. 205-237, 2003:

[8] C. Jonker, V. Robu, J. Treur. An Agent Architecture for Multi-Attribute Negotiation Using Incomplete Preference Information, In: Journal of Autonomous Agents and Multi-Agent Systems, 15(2), pp. 221-252, 2007.

[9] N. Matos, C. Sierra, N.R. Jennings. Determining Successful Negotiation Strategies: An Evolutionary Approach. In: Proc. of the Third Int. Conference on Multi Agent Systems, pp.182-189, 1998.

[10] J. Nadler, L. Thompson, L. van Boven. Learning Negotiation Skills: Four Models of Knowledge Creation and Transfer. In: *Journal of Management Science*, 49(4), pp. 529-540, 2003.

[11] H. Raiffa. *The Art and Science of Negotiation*, Harvard University Press, 1982.

[12] D. Zeng and K. Sycara. Bayesian Learning in Negotiation, In: *Int. Journal of Human Computer Systems*, 48, pp. 125-141, 1998.