Automatic Negotiation: Playing the Domain Instead of the Opponent

(Received 00 Month 20XX; final version received 00 Month 20XX)

An automated negotiator is an intelligent agent whose task is to reach the best possible agreement. We explore a novel approach to developing a negotiation strategy, a “domain-based approach”. Specifically, we use two domain parameters, reservation value and discount factor, to cluster the domain into different regions, in each of which we employ a heuristic strategy based on the notions of temporal flexibility and bargaining strength.

Following the presentation of our cognitive and formal models, we show in an extensive experimental study that an agent based on that approach wins against the top agents of the automated negotiation competition of 2012 and 2013, and attained the second place in 2014.

Keywords: Bargaining; Automated Negotiations; ANAC competition

1. Introduction

The problem of automated negotiation is one of the core research problems in Artificial Intelligence (Fatima, Kraus, & Wooldridge, 2014). Simply put, an automated negotiator is an intelligent agent whose task is to reach an agreement that will be best to his cause. This simple description hides many complexities as negotiation scenarios might differ in the parties incentives, available information, time constraints and others. Automated negotiation agents has many practical implications in many fields including electronic commerce (Mu-kun, 2010), resource management (Sim, 2006), mediation (Chalamish & Kraus, 2007), and task allocation (Lai, Li, Sycara, & Giampapa, 2004).

The negotiation problem has its roots in the economic literature, where it is often called the Bargaining problem (Binmore & Dasgupta, 1987). The appealing mathematical framework, built upon the utility theory, resulted in many solution concepts such as equilibriums, optimality principles, social welfare and others (Fatima et al., 2014). However, in complex negotiation scenarios, the economic theory quickly reaches its computational limitation and automated negotiators usually resolve to heuristics (Jonker, Robu, & Treur, 2007) or machine learning techniques (Oshrat, Lin, & Kraus, 2009). In addition to the economic literature, a bulk of research on the topic was conducted in the social and behavioral sciences, where the focus was on human behavior in different negotiation scenarios (e.g. gender differences (Katz, Amichai-Hamburger, Manisterski, & Kraus, 2008), cultures (Haim, Kraus, & Blumberg, 2010), and personalities (Zhang & Qiu, 2005)).

In order to advance practical research and application of automated negotiation agents, the “Automated Negotiation Agents Competition” (ANAC) was founded in 2010.¹ This annual competition allows researchers to employ automated negotiating agents, each carrying out a negotiation strategy, in order to carry out bilateral negotiation in non-trivial settings, and evaluate their relative performance. The evaluation is made in a

---

¹Website available and maintained at TUDelft.
round-robin fashion where each agent is pitted against each of the other agents, and the performance scores in each round are aggregated.

Examining the agents over the years (through data available on the competition website) we can see a strong shift in techniques from economic based solutions with heuristics, towards opponent modeling with machine learning techniques.\(^2\) It is well known that when interactions are repeated, learning the opponent and developing reciprocity become of prominent importance (Axelrod, 1984; Cheng, Zuckerman, Nau, & Golbeck, 2011). However, when the interactions are a series of one-shot bi-lateral negotiation encounters, we feel that focusing on opponent modeling should not be the corner-stone of constructing a capable automated negotiator. With that in mind, our research was motivated by the need to design automated negotiation strategies for complex negotiation domains, without any form of opponent modeling component.

For that we explore a different approach to developing negotiation strategy, which we call a **domain-based approach**. We use just two domain parameters, reservation value and discount factor, to cluster the negotiation domain to 4 different regions, in each of which we employ a predefined heuristic strategy which resembles a cognitive behavioral intuition based on the notions of **temporal flexibility** and **bargaining strength**.

To illustrate our approach we constructed an automated negotiation agent, called DoNA\(^3\) that is an implementation instance of the presented domain-based approach. We conducted an extensive empirical evaluation that shows that our agent outperforms the agents that won first place in both the 2012 and 2013 version of the Finals round in the ANAC competition.\(^4\) Following its success we enrolled DoNA to the 2014 ANAC,\(^5\) which introduced both very large negotiation domains and non-linear utility functions, and won the 2\(^{nd}\) place (out of 21 participants) in the individual utility category.\(^6\)

The main contributions of our research is as follows: (1) the development of the domain-based negotiation strategy, whereas its generality and simplicity is in complete contrast to the current trends in the negotiation agents community. (2) an extensive experimental study of the idea via the implementation of a highly successful domain-based negotiation agent.

Our insights contribute to the idea that playing the domain is at least as important as learning the opponent. It is important to note that while the notions of reservation values and discount factors have been explored in various contexts, for example to compute the equilibrium points of negotiation domains (Binmore & Dasgupta, 1987). To the best of our knowledge, this is the first agent that uses only these two parameters to define its strategy.

The paper is organized as follows: in Section 2 we formally present the negotiation problem at hand while Section 3 presents the abstract theoretical model of the problem. Section 4 describes the domain based cognitive model, and Section 5 discusses the implementation details of our agent. In Section 6 to present the extensive experimental results. In Section 7 we conduct a (limited) exploration of the sensitivity of some of the agent parameters. Section 8 discusses related work and Section 9 concludes the paper.

---

\(^2\) More than 90% of all agents in the last two competitions used some form of opponent model.

\(^3\) Available at \url{http://64.79.220.195/dona}

\(^4\) Moreover, it is important to note that in the 2013 competition the organizers allowed agents to conduct offline-learning between negotiation sessions by inputting a file of the negotiation session history to their agents. This information is of crucial importance as it can be used to learn the domains and the opponent strategies. However, our agent managed to win first place without using any information from previous sessions.

\(^5\) Took place in France in conjunction with the Autonomous Agents and Multiagent Systems conference (AA-MAS’14).

\(^6\) Second to Agent M of Nagoya Institute of Technology, however the difference in their score was not statistically significant.
and presents some research future directions.

2. Automated Negotiation

We study a bilateral, multi-issue negotiation setting in which two agents negotiate to reach an agreement on a set of several issues. There is a set of issues, denoted by $I$. Each issue, $i \in I$, has a finite set of possible values, denoted $O_i$. An offer, is denoted by \( o = \{o_1, o_2, \ldots, o_{|I|}\} \) where $O$ is the finite set of values for all issues, and $\forall i : 1 \leq i \leq |I|, o_i \in O_i$.

The negotiating parties’ utility functions, denoted $U_j(\bar{o})$, where $j$ denotes the player’s name, are time sensitive as they are discounted as a function of time. Each agent is assigned a discount factor, denoted $\delta_j$, which influences its utility as time passes. An agent’s discounted utility, denoted $U_j(\bar{o}, t)$ is calculated as follows:

$$U_j(\bar{o}, t) = U_j(\bar{o}) \times \delta_j^t$$

Where $U_j(\bar{o})$, $t$, and $\delta_j$ are real numbers normalized to the range $[0, 1]$.

The negotiation protocol is turn-based, similar to Rubinstein’s alternating offers protocol (Osborne & Rubinstein, 1994). Each side, in turn, can propose a possible agreement, accept the agreement offered by the opponent at the previous turn, or opt-out of the negotiation.

The negotiation can end either when: (a) the negotiators reach an agreement, (b) one of the agents opts-out, thus forcing the termination of the negotiation in disagreement, or (c) a pre-specified time limit is reached before an agreement is made. In case (a), each agent achieves a value depending on the agreement reached. In both cases (b) and (c), both agents achieve their reservation values. In all cases, the achieved value is discounted by the cost of time.

Last, we consider environments with incomplete information. That is, while both sides have a preference profile, each agent is only aware of its own profile and does not have any information about the profile of the other player as well as its discount and reservation values. The preference profiles are linearly additive utility functions (however, in the 2014 competition this assumption was relaxed and general utility functions were assigned).

3. An Abstract Theoretical Model

Our bilateral bargaining game consists of two agents, namely agent DoNA and agent Opponent, denoted Opp. The utility function for each agent is a private knowledge. Note that many bargaining games, e.g. (Kalai & Smorodinsky, 1975; Nash, 1950), assume that the parties’ utilities, and alternatives in case no agreement is reached, are common-knowledge, and there is full information. While we know $U_{DoNA}(\bar{o})$, $\delta_{DoNA}$, and the disagreement outcome, denoted $DisOutcome_{DoNA}$, we will need to make some assumptions regarding the $U_{Opp}(\bar{o})$ or regarding Opp’s strategy profile. As we later discuss, we choose the former, and make the assumption that the opponent holds a utility similar to ours. This assumption implies that any concession suggested by DoNA as part of the bargaining process should be received as such by the opponent, and essentially results in an increased $U_{Opp}(\bar{o})$.

---

7This models situations where one side is more constrained with respect to the time of agreement.

8For further discussion on this topic we refer the reader to (Lai & Sycara, 2009) and references there-in.
Having specified the agents participating in the bargaining game, and discussing our beliefs and knowledge of their utility, we complete the presentation of the game by discussing the actions available to the agents. At any point in time, each agent can serve as an offer-maker, suggesting an agreement offer $\vec{o}$ or as an offer-receiver deciding whether to accept or reject the offer. As the agents are assumed to be rational it should be rather intuitive that the main objective of each agent is to maximize her final utility achieved from bargaining agreement or disagreement.

As discussed above, we choose to abstract away from opponent modeling. Therefore our model boils down to a rather straightforward optimization problem which we present below and solve for every decision epoch.

$$\vec{o}^t_{DoNA} = \arg \max_{x_t \in \Omega_t} U_{DoNA}(x_t) * \delta^t_{DoNA}$$ \hspace{1cm} (1)

Subject to:

$$\max_{x_t \in \Omega_t} U_{DoNA}(x_t) * \delta^t_{DoNA} \geq \text{DisOutcome}_{DoNA},$$ \hspace{1cm} (2)

$$t \leq T,$$ \hspace{1cm} (3)

$$\Omega_t \subseteq \Omega_{t-1}, \quad \Omega_0 = O.$$ \hspace{1cm} (4)

The above optimization yields an optimal offer, with respect to utility maximization for DoNA. If DoNA is making an offer in this epoch, this is the offer she will make - it may be the case that she makes several offers which yielding her the same utility. The first constraint (eq. (2)), is the individual rationality constraint which compares the maximal discounted utility and the disagreement outcome - DoNA will not negotiate an agreement if she can reach a higher utility level by disagreement.\(^{10}\) The second constraint (eq. (3)), ensures we adhere to the (arbitrary) time limit set on the negotiations. The final constraint allows for progression of bargaining towards an agreement. The rather general presentation (eq. (4)) does not provide any limitations on the manner in which the subset is constructed. It could have replaced by a different constraint, e.g., $U_{DoNA}(\vec{o}^t_{DoNA}) \geq U_{DoNA}(\vec{o}^{t+\epsilon}_{DoNA})$ where $\epsilon \geq 0$ - the intuition for this is that in order to reach an agreement DoNA should make a concession, i.e. in the next time epoch her maximum achievable utility is lower than the maximum achievable in this time epoch - however, we chose not to include this in the model, keeping the model as simple and general as possible while maintaining the properties we wish to highlight. In order to solve for the optimal we determine the necessary first order condition (FOC) by deriving the discounted utility with respect to time and by noticing that this simple derivative is non-positive (as $\ln(\delta_{DoNA}) < 0$), i.e. the maximum utility level achievable by DoNA decrease over time.

$$\frac{dU_j(\vec{o}, t)}{dt} \leq 0.$$ \hspace{1cm} (5)

\(^9\)We focus our attention on an offering agent, as the decision to reject or accept are simple to conclude if we understand the strategy employed by an offering agent.

\(^{10}\)We note that in its current form the maximization is found with respect to a non-discounted disagreement outcome. This is purely a modeling choice which can be adjusted by multiplying the right hand side of eq. (2) by the scalar $\delta^t_{j}$. 

4
Figure 1. The quadrants of the domain-based strategic model

The FOC, coupled with our model constraints enables us to discuss when trying to reach an agreement should be attempted by the DoNA, as illustrated in Figure 1. Specifically, it should be clear that if the Individual Rationality constraint is breached, DoNA should not attempt to reach an agreement. This can happen in two cases distinguishable by parameter values. The first case is that of a “low” discount factor which is detrimental to the utility. In this case, denoted as Zone I and Zone II in Figure 1, the discount factor is so low that if an agreement is not reached immediately, agents should turn to their outside options for a higher terminal utility. The second case relates to lucrative outside options available to our agent. In this case, denoted as Zone II and Zone III in Figure 1, DoNA can achieve a high utility in case an agreement is not reached. The importance of this result is the understanding that if one (or both) of the cases occurs, negotiations to reach an agreement should not be carried out. Furthermore, it is also the insight that in all other cases, there should be an attempt to reach an agreement.

Figure 1 illustrates that given a set of parameters we can easily identify which of the four zones we occupy, and hence we should be able to conclude what approach to bargaining our agent should apply. For parameter values that are in Zone I, where time depreciates our maximal possible utility at a high rate, negotiations are expected to terminate quickly (the left hand side of the individual rationality constraint will be diminishing quickly). For parameter values in Zone III, our disagreement outcome is such that the right hand side of the individual rationality constraint is very high, resulting in forgoing negotiations and opting for the high utility provided by the outside option. For parameter values in Zone II the aforementioned effects are amplified as both occur simultaneously, i.e. low $\delta_{\text{DoNA}}$ and high reservation value. In Zone IV the effects are such that the possible utility achievable is higher than the outside option, even when accounting for the effect of discounting over time.

\footnote{We are aware that “Low”, “High” and “Higher” are relative terms. This issue is addressed and discussed in the empirical section of this paper.}
Next, we complement our initial model analysis by an additional discussion of the intuition behind the relative conditions presented above.

4. The Cognitive Model - Playing the Domain

In the negotiation problem presented above there are three aspects of uncertainty with respect to our opponent: opponent’s preferences, opponent’s reservation value, and opponent’s discount factor. As we do not want to explicitly model our opponent, we will not try to estimate the opponent’s utility function, but design a strategy based solely on our domains reservation and discount values.

Figure 1 presents a rough division of the set of possible domains based solely on the reservation value (x-axis) and the discount factor (y-axis). Looking at the relationships between the opposing corners of the graph we can see the following. The red diagonal, connecting the Play (Zone IV) and the End (Zone II) quadrants, signifies the temporal flexibility in the negotiation: In case that the discount factor is low and the reservation value is high (the “End” quadrant, i.e., Zone II), we wish to spend as little time as possible before either reaching an agreement or opting-out of negotiations, and therefore, this will be a lower bound on the amount of time we should spend on negotiation, implying a strategy where we opt-out of negotiations immediately. On the other edge of the time spectrum is the case where the discount factor is high and the reservation low (the “Play” quadrant, i.e., Zone IV), where we want to maximize the time allotted for negotiations in-order to enable an agreement. This can be thought of as an upper bound with respect to how much time to spend negotiating, implying we should allow for a ”step-by-step” concessions strategy - e.g., the Zeuthen strategy for the Monotonic Concession Protocol (Rosenschein & Zlotkin, 1994), which in the perfect information case has been shown to maximize the Nash Product if carried out by both negotiating parties - which should result, eventually in an agreement. In between these two bounds, some combination of these strategies should be used.

The green diagonal, connecting the Urgency (Zone I) and Restraint (Zone III) quadrants, provides a notion of our bargaining strength in the negotiation; In contrast, when we have a high reservation value and a high discount factor (i.e., low cost of negotiation process) we are in a favorable (strong) position and can afford being restrained, i.e., not make any concessions in order to achieve agreement, ultimately achieving agreement only in case the other side makes (what is perceived to be very-serious) concessions. On the other extreme, with low reservation value and a low discount factor, we are in an unfavorable (weak) position, and must make many concessions, and fast, to provide the other agent with the proper incentive to accept an agreement. This diagonal, the bargaining strength, should determine the concession strategy applied by the negotiating party.

As our suggested approach does not model the opponent (and to some extent do not even consider the opponent’s strategy) we base our analysis regarding the time allocated to the negotiations and our concession stance based on the given reservation value and discount factor. With the above interpretation in mind, the four quadrants provide us with cognitive-dominant behavior strategies.

**Play** — With a high discount factor (and therefore low cost of negotiation time) and a low reservation values the agent has a strong incentive to cooperate and achieve an agreement. As the cost of time is low, it has enough time to follow some concession protocol, thus participating in a classic interpretation of a negotiation game.
Urgency — With a low discount value (that is, high cost of negotiation time) and a low reservation value, the agent reasons that reaching an agreement is better than disagreement, and due to the high rate of depreciation doing it should be done as quickly as possible. In this case the strategic behavior dictates that a fast application of the concession strategy should be taken. For instance, by increasing the amount of concession steps per round.

Restraint — With high discount factor (that is, the cost of negotiation time is low) and a high reservation value, the agent understands that there is room for some benefit from cooperation, albeit very limited. Hence a negotiation session, if one takes place, is restraint, and it adopts a stubborn stances (i.e., hold out) and progress very slowly by, for example, reducing the number of concessions per round.

End — With a low discount factor and a high reservation value, the agent wants to end negotiation as fast as possible. This can be achieved by opting-out, or by reaching an agreement if one is agreed upon at the early stages of the negotiation, as time is very costly here.

The above behavioral quadrants span the entire space. While the behaviors are simple and explicit in the end points of the diagonals, any realization of the model should be constructed based on some procedure that takes into consideration the weighted distance of the quadrants strategies from the given reservation and discount values. Accordingly, each diagonal has two points: a minimum and a maximum, at the corners. These four points represent the natural behaviors when the reservation and discount values approach these respective limits.

5. DoNA - Implementation Details

Based on the behavioral model, we present the automated negotiation agent we have developed. We used the ANAC 2012 data to calibrate our model parameters, e.g., reservation value and discount factors thresholds to determine the agent’s negotiation strategy, time span to carry out negotiations, size of concession steps, etc. We use the ANAC 2013 as a test set for our model, and enrolled it to the ANAC 2014 to evaluate performance in a real-time competition (results are presented and discussed thoroughly in Section 6).

Focusing our attention on the values of the discount factor and the reservation, we discussed the possible behavioral strategies that our agent should be governed by when these values are at one of the 4 corners of the domain (see Figure 1). As mentioned between these points a combination of strategies should be used. Therefore, we found it reasonable to partition our state space into a 3-by-3 grid of combination of values, a division which allows us to be somewhat more accurate than a simple 2-by-2 grid yet allows for some differentiation between combinations of intermediate values of discount factors and reservation outcomes.

Table 1 describes our heuristics that were used in the nine regions of the reservation value (denoted by R) and discount factor (denoted by D) space after normalizing their values to the [0, 1] range. For both parameters, we defined the “low” range as (0, 0.25], the “medium” range as (0.25, 0.75], and the “high” range as [0.75, 1].

The following heuristics try to mimic the behavioral strategy described above in order to exemplify an instance of our cognitive model. We do not claim that the suggested

---

12Admittedly, we present what we refer to as satisfactory - rather than optimal - parameter values in the sense that the values we used perform well enough.
thresholds and heuristics are optimal. For instance, the above thresholds were chosen based on the parameters used in the final round of ANAC 2012; they were not fine-tuned, and it might be the case that performances could be further enhanced by learning and optimizing these threshold values. In addition, as we decided to shy away from opponent modeling techniques, we implicitly assume that our opponent holds similar discount and reservation values as our own (which was the case in all previous ANAC competition). Nevertheless, as we shall see in Section 6, DoNA’s excellent performance was achieved even when using this (sub-optimal) instantiation of the cognitive model.

For high reservation option range (right column), when the discount is medium or low, we use the \textit{End} heuristic, which ends the negotiation as soon as possible. The rationale for this is that we don’t have much to gain from negotiation, but if we wait, we forgo the high reservation value due to discounting over time.

At the low discount factor range (top row) when the reservation value is medium or low, we use the \textit{Fast} heuristic, which makes fast concessions in order to reach an agreement \textit{ASAP}. In this range, the agent considers all bids in the domain for which its utility is at least as high as the reservation value, and offers them to the opponent in decreasing order of utility, until one of the offers is accepted, or time expires, or the opponent makes a bid with utility higher than the next bid to be offered. The rationale is that the cost of time is high, so we strive to reach an agreement fast even if it means that we make a large concession.

For low reservation value and medium discount factor range (left-medium cell), we use the \textit{Reciprocal} heuristic, which makes a concession step towards the opponent whenever the opponent makes a concession to us.\footnote{We do not differentiate between concession steps of different size. This is indeed a place that can probably further optimize our performances.} To implement this heuristic, we keep two lists: a list of all bids in the domain ordered by decreasing order of utility for us, and a list of all bids sent by the opponent. Whenever the opponent offers a new bid (that does not already exist in our list), we send a new bid with the next-highest utility from our list. The rationale is that this cell is an average between the Fast strategy (at the top row), which makes many concession steps without waiting for mutual concessions, and the Last strategy (at the bottom row), which makes only few concessions and at the last stages of the negotiation.

At the high discount factor range (bottom row), we use the \textit{Last Moment} heuristic, which waits for the last moment to provide incentives (in form of concessions to the opponent) to try and strike an agreement. The rationale is that we have nothing to lose by waiting because the cost of time is low, and we might gain if the opponent decides to concede earlier. The Last Moment heuristic stalls the negotiation (by repeatedly offering the maximum utility attainable and not offering any concessions) until 75% of the time has elapsed. Meanwhile, we estimate the average time, $t_r$, required for a single round in which each agent makes a single negotiation offer. The “last moment” is defined as $t_r$ before timeout. After 75% of the time has passed, we start making concessions steps in a logarithmic fashion, as follows. Before half the time to timeout has passed (i.e. between 75% and 87.5% of the time has passed), we send all the bids with a utility of at least...
0.98 of the maximum attainable value. In the next logarithmic step (i.e. between 87.5% and 93.75% of the time has passed), we send all the bids with a utility of at least 0.96 of the maximum. We proceed in the same way, lowering our threshold in 0.02 utility units each logarithmic time step, until there are 2 rounds to time-out (i.e. the time is at least timeout minus \(2 \cdot t_r\)). At this point, we select between accepting the opponent’s offer, opting-out, or offering the bid that is best for us among all bids previously made by the opponent. At the last round (when the time is timeout minus \(t_r\)), we select between accepting the opponent’s offer and opting out.

Last, at the middle cell, we use a **Combined** heuristic, which uses an additional parameter, the **fairness equilibrium**. The fairness equilibrium, based on the **Zeuthen strategy** for the Monotonic Concession Protocol (MCP) (Rosenschein & Zlotkin, 1994), is the offer that is agreed upon if each of the two agents takes minimal concession steps from its respective maximum utility values until an agreement is reached. This equilibrium is “fair” in the common cultural interpretation as each side takes a similar number of concession steps until the middle point is reached. This occurs when one agent proposes an agreement that is at least as good for the other agent as their own proposal:

\[
\text{util}_{\text{fair}} = (u_1(x_2) \geq u_1(x_1)) \lor (u_2(x_1) \geq u_2(x_2))
\]

**Algorithm 1**

\[
\text{Algorithm 1 } \text{util}_{\text{fair}} (U_1, U_2)
\]

**repeat**

\[
x_i \leftarrow \text{arg max}_{x_i}(U_1(x_i))
\]

\[
x_j \leftarrow \text{arg max}_{x_j}(U_2(x_j))
\]

\[
\text{Delete}(x_i, x_j)
\]

**until** \(U_1(x_j) \geq U_1(x_i) \lor U_2(x_i) \geq U_2(x_j)\)

The algorithm runs as the agent takes concessions steps until the stopping criteria in true. When the algorithm stops, we look at the last offer presented and can calculate the percentage of utility we get with respect to the maximum attainable utility. In preliminary research, we have conducted an analysis of the domains previously used in the ANAC competitions and computed (using Algorithm 1) a cross-domain average \(\text{util}_{\text{fair}}\) value of 77%. This value is almost identical from both players perspectives (max difference < 0.03 between the sides). The algorithm works as follows: \(x_i\) refers to the offers made by agent \(i\), while \(x_j\) refers to the offers made by agent \(j\). The algorithms iterates over all the possible offers of the players, and each round it deletes the highest utility offer for each player. It continues removing offers until the utility that agent \(i\) or \(j\) get from the opponents’ next offer is higher than the utility that it gets from his own current offer.

With this parameter we can now present our **Combined** heuristic which uses the time dimension to mix between the aforementioned strategies and does so as follows.

1. Define \(T_{\text{end}}\) as the time in which we are going to end the negotiation. This time is calculated as the time in which the discount is exactly 0.8, i.e., \(\delta_{\text{DoNA}}^T_\text{end} = 0.8\).
2. During time period \([0, 0.25 \cdot T_{\text{end}}]\) use the Fast heuristic using a reservation value of \(\text{util}_{\text{fair}} \times 120\%\).
3. During time period \([0.25 \cdot T_{\text{end}}, 0.5 \cdot T_{\text{end}}]\) use the Fast heuristic using a reservation value of \(\text{util}_{\text{fair}} \times 110\%\).
4. During time period \([0.5 \cdot T_{\text{end}}, 0.75 \cdot T_{\text{end}}]\) use the Reciprocal heuristic.
5. During time period \([0.75 \cdot T_{\text{end}}, T_{\text{end}}]\) use the Last heuristic.

9
The rationale behind the order in the combined heuristic is that it defines the allotted time for negotiations. During the first part of this time a high utility may be achieved by progressing fast towards an agreement, when we restrict ourselves to high utility offers and note that we want to avoid losses due to the cost of time. If our fast concessions do not result in an agreement, we slow down and limit the pace of concessions to match those of the opponent using the Reciprocal heuristic, which is deemed more fair, and intuitively provides a higher probability of agreement. At the last period of the negotiation, we have formed a belief that our opponent is not willing to compromise, we allow for the minute chance he may concede but avoid any unnecessary concessions on our part, all the while guaranteeing achieving at least our discounted reservation value. As we have defined the 100% to be our minimal fair utility that we want to achieve, the values of 120% and 110% were selected ad-hoc as some intermediate utility that represent 0.95 and 0.85 of that value.

6. Experimental Evaluation

We have conducted extensive empirical evaluation by simulating the ANAC 2012 and ANAC 2013 final competitions. We let the DoNA agent play with the finalist agents in each competition, in all the negotiation domains from the final stage of the competitions (the agent code and domain data are publicly available from the competition website). After that, we have also enrolled our agent to the ANAC 2014 competition to evaluate him in real experiment.

6.1. Competition Details

Our evaluation were conducted in a similar fashion to the ANAC competitions final stage. The following procedure presents the round-robin competition:

(1) For each negotiation domain $Dom$, which includes two preferences profiles each with its respective $R$ and $D$:

(a) For each ordered pair of agents, $(a_i, a_j)$, do the following $K$ times, where $K$ is a given constant:

(i) Let agent $a_i$ negotiate with agent $a_j$, such that agent $a_i$ is one side of the domain $Dom$, and agent $a_j$ is the second side. The time limit for each negotiation were 3 minutes, similar to the original competition.

(ii) When the negotiation ends, give each agent a score determined by the result. I.e., if the negotiation ends with an agreement, give each agent a score calculated by summing its utilities for the values in the agreement, and if the negotiation ends without an agreement (either in case of opt-out or time-out), give each agent its reservation value.

(iii) In both cases, the score is discounted by the time discount factor as explained in the Negotiation Problem section.

(b) Calculate for each agent its average utility for the current $Dom$.

(2) For each agent $a_i$:

(a) Calculate its average score over all domains.
<table>
<thead>
<tr>
<th>Agent</th>
<th>Average score ↓</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>DoNA</td>
<td>0.587</td>
<td>0.007455</td>
</tr>
<tr>
<td>CUHKAgent</td>
<td>0.575</td>
<td>0.008011</td>
</tr>
<tr>
<td>OMGCagent</td>
<td>0.571</td>
<td>0.007267</td>
</tr>
<tr>
<td>AgentLG</td>
<td>0.557</td>
<td>0.008051</td>
</tr>
<tr>
<td>Meta-Agent</td>
<td>0.546</td>
<td>0.007164</td>
</tr>
<tr>
<td>IAMhaggler2012</td>
<td>0.540</td>
<td>0.006930</td>
</tr>
<tr>
<td>AgentMR</td>
<td>0.438</td>
<td>0.007114</td>
</tr>
<tr>
<td>BRAMAgent2</td>
<td>0.438</td>
<td>0.009240</td>
</tr>
</tbody>
</table>

6.2. **ANAC 2012 - development set**

Our first experiment was conducted similarly to the ANAC 2012 competition. We pitted DoNA against seven agents that passed the qualifying round and got to the final round.\(^{14}\) Here are the description of the final round agents.

**AgentLG** (Bar-Ilan University, 2nd place) — The agent uses time-dependent heuristics to offer bids in decreasing order down to a threshold based on the discount factor. It then decreases the threshold based on the estimated opponent’s utility profile. Finally, at the last moment, it offers the opponent’s best bid if it is higher than the reservation value.

**AgentMR** (Nagoya Institute of Technology) — The agent tries to estimate whether the domain is a win-win or a win-lose based on the utility from the opponent’s first bid, which is assumed to be its best bid. Based on this estimation, it calculates the concession rate. It makes bids that are similar to opponent’s bids in order to increase the probability of acceptance.

**BRAMAgent2** (Deutsche Telekom and Ben-Gurion University) — The agent seeks a win-win agreement by constructing a bid from the most required values from the opponent’s set of recent offers. It also tries to interrupt modeling attempts by the other agent. It ends the negotiation if the next candidate bid has a utility lower than the reservation value.

**CUHKAgent** (Chinese University of Hong Kong, 1st place) — The agent’s bidding strategy is based on selecting the estimated best bid for the opponent, from a set of negotiation outcomes that it finds acceptable. Its acceptance strategy is based on a non-exploitation threshold that is a function of the discount factor, domain size and concession degree of the opponent. Its termination strategy is based on treating the reservation value as an alternative offer proposed by the opponent.

**IAMhaggler2012** (University of Southampton) — This agent uses a decision theoretic approach to select the response that maximizes the expected utility, considering the discount factor, the remaining time and the future opponent concessions as predicted by a Gaussian process regression technique.

**Meta-Agent** (Ben-Gurion University) — This agent holds the set of publicly available ANAC 2011 agents, and selects and use the best agent for each domain, using a linear regression model.

**OMACagent** (Maastricht University, 3rd place) — This agent uses opponent modeling with discrete wavelet transform and cubic smoothing spline. It applies adaptive concession making using a function estimating the future opponent concessions.

According to the competition protocol in 2012, we use the domains in the ANAC

---

\(^{14}\) Codes can be found in http://anac2012.ecs.soton.ac.uk/ Due to technical reasons, we could not load the agent called “TheNegotiator Reloaded”, and one of the 24 final round domains.
competition where each domain was duplicated three times with different reservation and discount values. The reservation values were in \{0, 0.25, 0.5\} and the discount values were in \{0.5, 0.75, 1\}. Totally there were 69 different domains with varying number of possible outcomes (between 3 and 390,625). In each domain, each ordered pair of agents negotiated \( K = 5 \) times. The results are presented in Table 2.

The results show that DoNA agent achieved the top average score with 0.587, while the second place went to CUHKAgent (the winner in the original 2012 competition) with an average score of 0.575. The third place went to OMGCagent (who got third place in the original 2012) with 0.571 while AgentLG moved from being second in the original competition to the fourth place. The column titled CI denotes the 95% confidence interval of the results (in a student’s T-test). It is worth noting that the order of the agents in the original competition can be changed in the reruns or the competition with the introduction of new agents because of the aggregation of the session with the new agent.

### 6.3. ANAC 2013 - test set

The 2013 competition had a profound change in the rules, which allowed agents to learn between sessions. They have the ability to save information during and after negotiation session, and load it at the beginning of new session on the same domain and profile. This change allows the agents to learn between sessions, for example to model the opponent’s profile more accurately. The competition involved 19 teams from eight institutions. Following the qualifying round, a total of seven agents proceeded to the final round.\(^{15}\) Here are their description:

**AgentKF** (Tokyo University) — Based on the average and variance of opponent’s bids in our utility function, the agent classified the opponent as “uncooperative”, “neutral” or “cooperative”. This classification affects the acceptance threshold. Its bidding strategy involves searching for nearly-Pareto-optimal bids by selecting bids that are similar to both our preferred bids and the opponent’s bids.

**GAgent** (Nagoya Institute of Technology) — This agent tries to estimate whether the domain is win-win or win-lose based on our utility from the opponent’s first bid, which is assumed to be its best bid. Based on this estimation and on the remaining time, it calculates the acceptance threshold value. While offering bids, it makes concessions only if the opponent makes mutual concessions.

**Slava Agent** (Bar-Ilan University) — At the first 95\% of the time, it sends random bids with utility at least 0.95, hoping that the opponent will learn what we want and send us an acceptable offer. At the last 5\% of the time, it insists on the bid that is best for us, hoping that the opponent will concede.

**TMF-Agent** (Ben Gurion University) - 3rd place — it uses opponent modeling based on estimating the utility of each value for all issues. Its bidding strategy is to select a bid that maximizes the weighted average of our utility and the opponent’s utility, where the weight of each utility depends on a hardness parameter which is a function of time and discount factor. If the utility of this bid is lower than the reservation value, it ends immediately.

**Slinkhard** (Delft University) - 1st place — Its bidding strategy is based on predicting the opponent’s strategy and it also implies opponent modeling based on discrete wavelet prediction and frequency model.

\(^{15}\)http://www.itolab.nitech.ac.jp/ANAC2013/
Table 3. Results of rerun of ANAC 2013 with DoNA

<table>
<thead>
<tr>
<th>Agent</th>
<th>Average score</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>DoNA</td>
<td>0.637</td>
<td>0.010249</td>
</tr>
<tr>
<td>Meta-Agent</td>
<td>0.628</td>
<td>0.011084</td>
</tr>
<tr>
<td>InoxAgent</td>
<td>0.612</td>
<td>0.011704</td>
</tr>
<tr>
<td>Slinkhard</td>
<td>0.609</td>
<td>0.011056</td>
</tr>
<tr>
<td>TAgent</td>
<td>0.606</td>
<td>0.011949</td>
</tr>
<tr>
<td>TMF-Agent</td>
<td>0.591</td>
<td>0.011571</td>
</tr>
<tr>
<td>AgentKF</td>
<td>0.567</td>
<td>0.012841</td>
</tr>
<tr>
<td>Slava Agent</td>
<td>0.537</td>
<td>0.013605</td>
</tr>
</tbody>
</table>

Meta-Agent (Ben-Gurion University) - 2nd place — The newer version of this agent used the best agent for each domain from the set of 2012 ANAC agents, selected using the CART machine learning algorithm. It features include domain size, average utility of bids, discount factor, opponent first bid utility, average utility of relevant bids, and more.

Inox Agent (Delft University) — This agent’s bidding strategy is based on estimating the Kalai-point. It also uses opponent modeling based on a simple frequency model. An opponent bid is accepted if it is better than our upcoming bid AND not worse than the median of agreements from previous negotiations.

The distribution of the domains with respect to the reservation and discount values can be seen in Figure 2.16

The results are described in Table 3. In each domain, each ordered pair of agents negotiated $K = 10$ times. We can see that DoNA managed to win the first place before Meta-Agent (who was also ranked second in the original competition). They were followed by InoxAgent who managed to upgrade itself to the third place while Slinkhard, the original winner, grabbed the fourth place.

It is important to note at this point that new rules that were presented in 2013 dictate that agents can use learning in-between sessions, using the logs of the previous sessions. Therefore, in 2013 the construction of a learning algorithm both for the opponent model and the domain in play was very important to the success of the agents. Nevertheless, as our agent was constructed before ANAC 2013 took place, in the presented results DoNA competes **without using the data from previous sessions**.

Following the previous experiment on the 2013 competition we conducted further anal-

---

16Two of the domains had different reservation values and discount factors for the players.
Table 4. Results of Finals ANAC 2013 agents on the new domains

<table>
<thead>
<tr>
<th>Agent</th>
<th>((R,D))</th>
<th>((0.75,0.25))</th>
<th>((0.95,0.05))</th>
<th>((0.99,0.01))</th>
</tr>
</thead>
<tbody>
<tr>
<td>DoNA</td>
<td>0.726</td>
<td>0.884</td>
<td>0.892</td>
<td></td>
</tr>
<tr>
<td>Meta-Agent</td>
<td>0.584</td>
<td>0.793</td>
<td>0.726</td>
<td></td>
</tr>
<tr>
<td>TMF-Agent</td>
<td>0.337</td>
<td>0.296</td>
<td>0.274</td>
<td></td>
</tr>
<tr>
<td>InoxAgent</td>
<td>0.326</td>
<td>0.293</td>
<td>0.246</td>
<td></td>
</tr>
<tr>
<td>Slinkard</td>
<td>0.329</td>
<td>0.290</td>
<td>0.265</td>
<td></td>
</tr>
<tr>
<td>GAgent</td>
<td>0.324</td>
<td>0.269</td>
<td>0.250</td>
<td></td>
</tr>
<tr>
<td>Slava Agent</td>
<td>0.324</td>
<td>0.263</td>
<td>0.210</td>
<td></td>
</tr>
<tr>
<td>AgentKF</td>
<td>0.206</td>
<td>0.221</td>
<td>0.106</td>
<td></td>
</tr>
</tbody>
</table>

Analysis of DoNA’s specific performances in the different domains. The analysis shows that while in most areas she achieved the first or second overall place, there are three domains which resulted in lower ranking (7th, 7th, and 8th). These domains are ones in which the discount was borderline, around 0.75 and both sides have the ability to get high utility values (a “win-win” domain). While DoNA waited for the last moment to get a favorable offers, the other agents, with the available information between sessions (the new rule for the 2013 competition) were able to get higher utilities faster at the early stages of the negotiations. We expect that with the availability of the between-session information, such weaknesses could be avoided. In addition, DoNA can use a smooth strategy at this zone.

As we saw in the above results (Table 3), the DoNA agent performs better than the other agents in the rerun of the 2013 competition despite the handicap on not using the information that is available between negotiation sessions. An addition handicap was the following. In the 2012 and 2013 competitions, each participating agent could suggest a single domain that will be part of the competitions’ domains. This, in theory, would provide a round in which the “hosting” agent (the one who suggested the domain) would most probably dominate the other agents. In the presented results (Tables 2 and 3) we did not include such domain for the DoNA agent.

With the ability to present a domain, like the other participating agents, we experimented with three domains that are centered around the upper-right corner of the cognitive model, an area that, apparently, was not very popular in previous competitions. Table 3 shows the results of the three additional domains with reservation values and discount factors \((R,D)\) set to the following values: \((0.75,0.25)\), \((0.95,0.05)\), \((0.99,0.01)\). These values sample different sections in the upper-right square in the model, that were absent from the finals of the 2012 and 2013 competitions (as can be seen in Figure 2).

From the results we can see that the DoNA agent, as expected, is the most dominant agent by a large margin (the results are significant in the 95% interval and are ordered according to the middle column). Also, we can see that our advantage will grow as close as we can get to the \((1,0)\) limits. We also were very surprised to learn that aside for Meta-Agent, which itself does not have a unique strategy, all the other agents where less successful with a difference of at least 0.39, 0.59, and 0.62 respectively. This exemplifies their general weakness in domains with these characteristics. It is easy to see that if this domain would be integrated in the competition, DoNA’s results would have been even better.

6.4. **ANAC 2014 - real competition**

After developing the agent on the 2012 data and testing it offline on the 2013 data, it was time to put it to the test. As such, we have enrolled DoNA to the 2014 ANAC
competition. However, as in previous years, the 2014 edition of the competition issued additional challenges: First, for the first time the utility functions of the domains were extended to include nonlinear utility functions. Second, the organizers have decided to include also very large size domains (e.g. domains of size $10^{50}$). However, the rule from the 2013 competition in which the bid history was available for learning prior to the negotiation was removed from the 2014 competition. In other words, agents only have the opportunity to adapt and learn from the offers they receive within a single negotiation session.

The effects of the new changes meant that the players could query the framework for the utility values of individual bids and get a utility number. However, due to time constraints and very large domains it was only possible to get a very small sample of the available bids and their utilities prior to the beginning of the negotiation session (around 100,000 bids). These changes meant that brute-force methods are not feasible, and previous techniques that were applied to modeling the opponent are also not relevant anymore. Nevertheless, DoNA did not use neither of these methods, so no changes were needed with that respect. However, DoNA was explicitly using the ordering of the bids in its strategy (e.g. when taking step-by-step concessions), this cannot be done explicitly in the new setting. In order to cope with the new challenges, when we were faced with domains that contained more than one million bids we assumed the utility function has a normal distribution on the utility values, and sample 100,000 bids and use them to sort all the bids we have. We also calculate the average and the variance of the sampled bids in order to set a threshold representing the minimum expected utility (which is a required data for DoNA’s strategy). We did not do anything else with respect to the nonlinear utility functions challenge since it was irrelevant to our model. In smaller domains DoNA acted as described in the original model.

The competition included 12 nonlinear utility domains: 4 small domains ($10^{10}$), 4 medium domains ($10^{30}$), and 4 large domains ($10^{50}$). 21 teams entered the competition from 13 different institutions in 8 countries. In the qualification stage the agents were randomly assigned to 3 groups, and the 3 top teams in each group advanced to the final stage. The qualification group in which DoNA was assigned to and its results are

---

17 Competition website: http://www.itolab.nitech.ac.jp/ANAC2014/
18 It took two weeks to run the qualification stage on a cluster of 6 computers.
The results show that DoNA was the top agent in his group by a significant amount, and together with AgentM and BraveCat advanced to the final round.

The final round was played with 10 finalists, on 12 nonlinear domains and the results are averaged over 5 trials, each repeated one time to prohibit learning. DoNA won the second place in the competition ANAC 2014 after the first place "AgentM" which uses a genetic algorithm with in its strategy (though its specific details are yet to be published). It is interesting to note that both DoNA and AgentM were in the same qualification group together, where in it DoNA performed better than AgentM.

In addition to the goal of maximizing the agent’s individual utility, the official results also included information about the agents’ performance when maximizing the social utility (i.e. maximizing utility score for both agents). In the qualification group, DoNA achieved a social utility score of 1.285 which was ranked 2nd in its group, behind BraveCat who came 1st with 1.422 and before AgentM who came 3rd in the social utility score with 1.282. However, in the final round DoNA did not do well in the social welfare category. Out of 8 agents in the final round, DoNA attained the 7th place with a score of 1.473.

We believe that it is much harder to design a social welfare maximization strategy based only on the domain parameters. This requires looking into the opponent moves, modeling it, avoiding exploitations, dealing with issues of reputation and trust, which was out of scope for this research effort.

### 7. Analysis and Discussion

The main motivation of our research was to explore the notion of a negotiating strategy that is based on the domain parameters and not opponent modeling. While DoNA is a successful exemplar of our approach, its implementation is not optimal and we believe that further fine-tuning of its parameters would result in a better performing agent. In order to illustrate this prospect we have conducted an additional set of experiments to evaluate the sensitivity of our implementation to parameter values.
7.1. Grid threshold values

We start by looking at the threshold value of the heuristic grid. In our DoNA’s implementation we divided the range to “low”, “medium” and “high” values using the 0.25, 0.5 and 0.75 thresholds. This division is ad-hoc and we cannot guarantee that these values provide an optimal, or even good, assignment of domains to heuristics. Therefore, in the next experiment we tried varying the threshold of the “high” discount factor from its original value of 0.75 to different values: 0.8, 0.87, and 1. This means, for example, that domains with a 0.77 discount factor and a value of 0.8 as its reservation value, will now play the “End” heuristic, while in the original implementation it used the “Last” heuristics.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>0.528916</td>
</tr>
<tr>
<td>0.8</td>
<td>0.51508</td>
</tr>
<tr>
<td>0.87</td>
<td>0.570874</td>
</tr>
<tr>
<td>1</td>
<td>0.541825</td>
</tr>
</tbody>
</table>

From the 2012 competition, we have isolated 6 domains that are relevant to the change in that threshold. These are the following domains: \{animal 0.75, camera 0.75, nice 0.75, outfit 0.75, kitchen 0.8, smart 0.87\}. We reran the competition on those domains and the results of DoNA’s performance on these domains are given in Table 7. We can see that DoNA’s original implementation (with 0.75 as its discount threshold) was not the optimal one. Using 0.87 as its high discount factor would have yielded a score of 0.570874 which is significantly better than its own score on these 6 domains (0.528916).

7.2. Steps sizes in “Last Moment” heuristic

A second experiment was conducted to explore the sensitivity of the logarithmic steps in the “Last Moment” heuristic. This function uses a log function to decide when to take the next concession step and to compute a new acceptance threshold in each step. This log-step behavior is similar in essence to the ideas presented in (Baarslag & Hindriks, 2013) where the authors view the acceptance strategy as a stochastic process and by clustering the opponents to different classes they can approximate their acceptance range.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.574781</td>
</tr>
<tr>
<td>2.0</td>
<td>0.615082</td>
</tr>
<tr>
<td>2.5</td>
<td>0.61171</td>
</tr>
</tbody>
</table>

Our original implementation uses log 2, which means that the agent considers a new concession step or an acceptance threshold in rounds 2 before the end of negotiations, then 4 rounds before the end of negotiations, 8, 16, and so on. We have conducted additional two experiments with log values of 1.5 and 2.5 and compared their behaviors on the 2012 version of the competition with all the domains. Looking at the results (Table 8) we can see that the best results were attained with the original value that was used in DoNA.
7.3. *DoNA* vs. *DoNA*

Another interesting question that comes to mind is how would *DoNA* fare against itself? Obviously, as *DoNA* is a domain-based agent, the result would depend on the domain parameters themselves. On high reservation values, when playing the “End” heuristic, when *DoNA* will play against itself they will both call for a quick conclusion to the negotiation and will most probably settle on the reservation value. In low discount values, when playing the “Fast” heuristic, both DoNASs will most probably attain higher values than any other agent because they will both push for a quick agreement to lose as little as possible due to time. On high values both in the reservation and discount factors, when playing the “Last” heuristic, both DoNASs will high social welfare but because of the delay (up to 75% of the time) it will not get the optimal possible value, and other agents can do better. When playing in the “Reciprocal” domains, *DoNA* against itself will play according to the fairness equilibrium, based on the Zeuthen strategy for the Monotonic Concession Protocol (MCP). This is basically the Nash solution of the game as shown in (Rosenschein & Zlotkin, 1994). In the “combined” square the agents will conclude on the same agreement as the “Fast” heuristic.

From the above analysis we can easily conclude that having another DoNA or two in the competition would benefit both of them. This is so because *DoNA* does not try to model and “outsmart” its opponent, but settle for a “good enough” result based on the domain in question. When the opponent also takes this approach, they would both benefit. To evaluate our hypothesis, we ran *DoNA* against itself on the domains in the 2012 competition and attained a value of $0.59937$. Comparing this to the average score of $0.572096$ that *DoNA* attained in the original competition against all agents, we can see that having an additional DoNA in the competition’s poll of agents, will improve their performances.

8. Related work

In this work we are looking at the problem of bilateral, incomplete information, time constrained, multi-issue negotiations. In other words, we have two parties that need to agree on several issues, but are unaware of the preferences of the other side. There is an extensive literature on different versions of the problem and solution approaches, we will cite the most relevant ones to point the differences with our current work. We refer the reader to an excellent review in (Lai & Sycara, 2009).

Game theoretic models provide a right mathematical tools to analyze and predict the negotiation outcomes. However, the cost of the mathematical rigidity is in the simplifying assumptions that are used. Their well-known equilibria solutions such as Nash bargaining, or the Kalai-Smorodinsky solutions (Kalai & Smorodinsky, 1975; Nash, 1950) requires complete information. In (Harsanyi & Selten, 1972) the authors provide a general solution for a bilateral bargaining game with incomplete information, however they look at single-issue negotiation without the notion of time discount. Game theoretic solutions provide many interesting insights on the formal side of the problem, but are usually not practical to apply in reality due to their simplifying assumptions.

Due to the inherent difficulty of the problem and the need for practical solutions, various heuristic models were proposed over the years. In (Byde, Yearworth, Chen, & Bartolini, 2003) the authors developed AutONA, an automated negotiation agent. Their problem domain involves multiple negotiations between buyers and sellers over the price and quantity of a given product. In (Jonker et al., 2007) the authors created an
agent to handle multi-attribute negotiations which involve incomplete information. The QOAgent (Lin, Kraus, Wilkenfeld, & Barry, 2008) is a domain independent agent that can negotiate with people in environments of finite horizon bilateral negotiations with incomplete information. The negotiations consider a finite set of multi-attribute issues and time-constraints. The KBAgent (Oshrat et al., 2009) is currently among the best human-agent negotiation agent. It also considers negotiations with a finite set of multi-attribute issues and time-constraints, and has been shown to be the most effective agent in achieving agreements with people in several domains involving multiple attributes. Recently the first step in migrating from Dialog-based environments to complete Chat-based interfaces was taken (Zuckerman, Rosenfeld, Kraus, & Segal-Halevi, 2013).

With respect to recent works that rely on the structure of the negotiation domain to convey appropriate negotiation strategies we can refer to (Sanchez-Anguix, Aydogan, Julian, & Jonker, 2014) in which the authors proposed a negotiation strategy for negotiation teams based on issue preferential compatibility and domain knowledge. They classify issues into those that are predictable and compatible among team members, and those that are unpredictable among team members. Then, they apply appropriate mechanisms to provide efficient results for the team. Additionally, in (Baarslag, Hindriks, & Jonker, 2014) the authors proposed several efficient bid acceptance mechanisms based on several domain factors, while in (Fujita, Ito, & Klein, 2014) they analyzed the negotiation domain at hand in order to find subspaces that are independent of each other. This way they tackle computational complexity and issue interdependence.

The usefulness of the different approaches have been recently put to the test in the yearly ANAC competition that started in 2010. The competition provides a platform to compare and benchmark different state-of-the-art heuristics developed for automated, complex bilateral negotiation. As space is limited we cannot describe all the approaches that were proposed, but here are a few notable examples: Yushu (2nd in 2010) uses learning based on the ten best recent proposals of the opponent, and estimates the number of rounds that are left for the negotiation to obtain the target utility. It also keeps track of the minimum utility it is willing to accept. OMACagent (3rd in 2012) stands for Opponent modeling and Adaptive concession-making. It uses discrete wavelet transform and Cubic smoothing spline to model the opponent utility as a function of time, and an estimation function for future opponent concession. Meta Agent (close 2nd in 2013) employs a machine learning algorithm selection method using ANAC 2012’s algorithms. Given a new domain, it compares all performance prediction of all agents on the domain and use the agent with the highest predicted utility. Nevertheless, none of the previous agents was basing its strategy only on the properties of the domain and its own disagreement utility and discount parameters.

In (Fatima, Wooldridge, & Jennings, 2004) the authors discuss the differences between issue-by-issue and a package deal negotiation. They use an incomplete-information setting, except for the reservation prices which is known and utilized for the concession list. In addition, their results are based on their fixed and constrained negotiation protocol. Other works address multi-issue negotiation where issues have binary values. For instance, in (Robu, Somefun, & La Poutré, 2005) an approach based on graph theory and probabilistic influence network is proposed to deal with multiple binary issues.

In other works opponent modeling procedures were develop to deal with the incomplete information attribute of the domain. For example, in (Li & Tesauro, 2003) the agent applies depth-limited search to find the most favorable offer based on its knowledge of the opponent’s type. Rejections are used to update the opponent model with Bayesian update rules. Nevertheless, the authors assume some knowledge about the agent’s utility function as defined by its type.
Another interesting aspect in negotiation is the notion of fairness. Examples from game-theory perspective can be see in (Raith, 2000), where a “fair” procedure based on the idea of auctions was introduced. First, all items are assigned to the one who values the items most, and then through monetary transfers, fairness is established. Another example can be found in (Brams & Taylor, 1996), were the “Adjusted Winner” procedure assigns to each agent the item it values most (not all items as above), and some money is being transferred to equalize their gains.

9. Conclusions

The current research trends in constructing automated negotiation agents is through the development of an accurate model of the opponent. Nevertheless, we conjecture that in cases where negotiations are a one-shot interaction, focusing on complex opponent modeling might not be the right path to take. Consequently, we present and study a domain-based negotiation approach, which utilizes a simplified cognitive model that defines the natural-heuristics which only accounts for the reservation and discount values of the domain. We provided an intuitive interpretation of these strategies in our agent, DoNA, and conduct extensive experimental works on data available from the 2012 and 2013 ANAC competitions. Following that, we enrolled DoNA to the 2014 ANAC.

Our results show that DoNA was at least as good as the other agents as it managed to grab the first place in an experimental rerun of the 2012 and 2013 competitions, without using any learning to utilize the available information between sessions (that was available in the 2013 competition). On top of that, DoNA won the 2nd place in the 2014 competition thus showing that it is also capable to dealing with non-linearity and large domain. These demonstrate, for the first time, that a simple domain-based agent can achieve results comparable to or even better than more complex opponent-based agents. We believe our agent can be used as a baseline for evaluating the utility of adding opponent-based models to agents.

In the future we see several possible directions. First, we conjure that DoNA will be especially successful when negotiating with people because it is built upon cognitive behavioral foundations for negotiation. We feel that humans would tend to apply the same reasoning process thus allow it to form an agreement faster. We plan on conducting an experimental session with students in the near future to evaluate DoNA’s performance when negotiating with real humans. Second, DoNA itself is hardly optimized with respect to its opponent in the competition, we plan to integrate a learning algorithm (based only on the domain’s parameters, not the opponent model) or find a continuous function for the thresholds and enlist DoNA to the coming ANAC competition. Finally, we are in the process of developing an algorithm to tune the parameters to different types of domains (“win-win” vs. “win-loss” vs. “loss-loss”), different sizes of domains (in terms of the number of possible agreements), and other possible parameters.

References


