

Automated Negotiating Agent with Strategy Adaptation for Multi-times Negotiations

Katsuhide Fujita

Abstract Bilateral multi-issue closed negotiation is an important class for real-life negotiations. Usually, negotiation problems have constraints such as a complex and unknown opponent's utility in real time, or time discounting. In the class of negotiation with some constraints, the effective automated negotiation agents can adjust their behavior depending on the characteristics of their opponents and negotiation scenarios. Recently, the attention of this study has focused on the interleaving learning with negotiation strategies from the past negotiation sessions. By analyzing the past negotiation sessions, agents can estimate the opponent's utility function based on exchanging bids. In this paper, we propose an automated agent that estimates the opponent's strategies based on the past negotiation sessions. Our agent tries to compromise to the estimated maximum utility of the opponent by the end of the negotiation. In addition, our agent can adjust the speed of compromise by judging the opponent's Thomas-Kilmann Conflict (TKI) Mode and search for the pareto frontier using past negotiation sessions. In the experiments, we demonstrate that our agent won the ANAC-2013 qualifying round regarding as the mean score of all negotiation sessions. We also demonstrate that the proposed agent has better outcomes and greater search technique for the pareto frontier than existing agents.

1 Introduction

Negotiation is an important process in forming alliances and reaching trade agreements. Research in the field of negotiation originates in various disciplines including economics, social science, game theory and artificial intelligence (e.g. [5, 6, 14, 16]). Automated agents can be used side-by-side with a human negotiator embarking on an important negotiation task. They can alleviate some of the effort required of people during negotiations and also assist people that are less qualified in the negotiation

K. Fujita (✉)

Faculty of Engineering, Tokyo University of Agriculture and Technology,
Tokyo 184-8588, Japan
e-mail: katfuji@cc.tuat.ac.jp

process. There may even be situations in which automated negotiators can replace the human negotiators. Another possibility is for people to use these agents as a training tool, prior to actually performing the task. Thus, success in developing an automated agent with negotiation capabilities has great advantages and implications.

Motivated by the challenges of bilateral negotiations between automated agents, the automated negotiating agents competition (ANAC) was organized [8, 10, 11, 22]. The purpose of the competition is to facilitate research in the area of bilateral multi-issue closed negotiation. The setup at ANAC is a realistic model including time discounting, closed negotiations, alternative offering protocol, and so on. By analyzing the results of ANAC, the stream of the strategies of automated negotiations and important factors for developing the competition have been shown [1]. Also, some effective automated negotiating agents have been proposed through the competitions [2, 12, 23].

Recently, for automated negotiation agents in bilateral multi-issue closed negotiation, attention has focused on interleaving learning with negotiation strategies from past negotiation sessions. By analyzing the past negotiation sessions, agents can adapt to domains over time and use them to negotiate better with future opponents. However, some outstanding issues regarding them remain, such as effective use of past negotiation sessions. In particular, the way of understanding the opponent's strategy and negotiation scenarios from the past sessions is unclear. In other words, it is still an open and interesting problem to design more efficient automated negotiation strategies against a variety of negotiating opponents in different negotiation domains by utilizing the past negotiation sessions.

In this paper, we propose an adaptive strategy based on the past negotiation sessions by adjusting the speed of compromising depending on the opponent's strategy, automatically. For judging the opponent's strategy, we need to characterize the opponents in terms of some global style, such as negotiation styles or a known conflict-handling style. One important style is the Thomas-Kilmann Conflict Mode Instrument (TKI) [13, 17]. The TKI is designed to measure a person's behavior in a conflict situation based on the concerns of two people appearing to be incompatible. The proposed agent tries to compromise speedily when the opponent is cooperative and passive. By employing this strategy, our agent achieves an agreement in the earlier stage compared with existing negotiating agents. If agents achieve an agreement in the earlier stage, agents can gain more utility because the time-discounted factor decreases the total utility. In addition, our agent has an effective search strategy for finding the pareto optimal bids.

In the experiments, we demonstrate that the proposed agent outperforms the other agents that participated in the qualifying round of ANAC-2013. We also compare the performance of our agent with that of the state-of-the-art negotiation agents. By analyzing the results, it is clear that our agent can obtain higher mean utilities against a variety of opponents in the earlier steps. Additionally, we demonstrate the change of the utility in multi-times negotiation for analyzing the learning strategies.

The remainder of the paper is organized as follows. First, we describe related works. Second, we show the negotiation environments and our proposed agent's basic strategy. Third, we propose a way of adjusting the compromising speed, and

a search method for finding pareto optimal bids. Then, we demonstrate the overall results of the qualifying round of ANAC-13 and some experimental analysis. Finally, we present our conclusions.

2 Related Works

This paper focuses on research in the area of bilateral multi-issue closed negotiation, which is an important class of real-life negotiations. Closed negotiation means that opponents do not reveal their preferences to each other. Negotiating agents designed using a heuristic approach require extensive evaluation, typically through simulations and empirical analysis, since it is usually impossible to predict precisely how the system and the constituent agents will behave in a wide variety of circumstances.

Motivated by the challenges of bilateral negotiations between people and automated agents, the automated negotiating agents competition (ANAC) was organized in 2010 [8, 10, 11, 22]. The purpose of the competition is to facilitate research in the area of bilateral multi-issue closed negotiation. The declared goals of the competition are (1) to encourage the design of practical negotiation agents that can proficiently negotiate against unknown opponents and in a variety of circumstances, (2) to provide a benchmark for objectively evaluating different negotiation strategies, (3) to explore different learning and adaptation strategies and opponent models, (4) to collect state-of-the-art negotiating agents and negotiation scenarios, and make them available to the wider research community. The competition was based on the GENIUS environment, which is a General Environment for Negotiation with Intelligent multi-purpose Usage Simulation [15].

By analyzing the results of ANAC, the stream of the strategies of ANAC and important factors for developing the competition have been shown. Baarslag et al. present an in-depth analysis and the key insights gained from ANAC 2011 [1]. This paper mainly analyzes the different strategies using classifications of agents with respect to their concession behavior against a set of standard benchmark strategies and empirical game theory (EGT) to investigate the robustness of the strategies. It also shows that the most adaptive negotiation strategies, while robust across different opponents, are not necessarily the ones that win the competition. Furthermore, our EGT analysis highlights the importance of considering metrics.

Chen and Weiss proposed a negotiation approach called OMAC, which learns an opponent's strategy in order to predict future utilities of counter-offers by means of discrete wavelet decomposition and cubic smoothing splines [3]. They also present a negotiation strategy called EMAR for this kind of environment that relies on a combination of Empirical Mode Decomposition (EMD) and Autoregressive Moving Average (ARMA) [4]. EMAR enables a negotiating agent to acquire an opponent model and to use this model for adjusting its target utility in real time on the basis of an adaptive concession-making mechanism. Hao and Leung proposed a negotiation strategy named ABiNeS, which was introduced for negotiations in complex environments [9]. ABiNeS adjusts the time to stop exploiting the negotiating partner and also

employs a reinforcement-learning approach to improve the acceptance probability of its proposals. Williams et al. proposed a novel negotiating agent based on Gaussian Processes in multi-issue automated negotiation against unknown opponents [23]. Fatima et al. focus on the bilateral multi-issue negotiation between self-interested agents in time-limitation settings [7]. By showing the negation model and the optimal procedure for each party, this paper determined equilibria for each procedure for two different information settings. Kawaguchi et al. proposed a strategy for compromising the estimated maximum value based on estimated maximum utility [12]. These papers have been important contributions for bilateral multi-issue closed negotiation; however, they don't deal with multi-times negotiation with learning and reusing the past negotiation sessions.

Recently, some studies have focused on the divided parts of negotiating strategies in the alternative offering protocol: proposals, responses, and opponent modeling. Effective strategies can be achieved by combinations of these strong strategies depending on the opponent's strategies and negotiation environments. Many of the sophisticated agent strategies that currently exist are comprised of a fixed set of modules. Therefore, the studies for proposing the negotiation strategies focusing on the modules are important and influential. Baarslag et al. focus on the acceptance dilemma: accepting the current offer may be suboptimal, as better offers may still be presented [2]. On the other hand, accepting too late may prevent an agreement from being reached, resulting in a break off with no gain for either party. This paper proposed new acceptance conditions and investigated correlations between the properties of the negotiation environment and the efficacy of acceptance conditions.

3 Negotiation Environments

3.1 *Bilateral Multi-issue Closed Negotiation*

The interaction between negotiating parties is regulated by a *negotiation protocol* that defines the rules of how and when proposals can be exchanged. The competition used the alternating-offers protocol for bilateral negotiation as proposed in [18, 19], in which the negotiating parties exchange offers in turns. The alternating-offers protocol conforms with our criterion to have simple rules. It is widely studied in the literature, both in game-theoretic and heuristic settings of negotiation [5, 6, 14, 16].

For example, *Agents A and B* take turns in the negotiation. One of the two agents is picked at random to start. When it is the turn of agent X (X being A or B), that agent is informed about the action taken by the opponent. In negotiation, the two parties take turns in selecting the next negotiation action. The possible actions are:

Accept: This action indicates that the agent accepts the opponent's last bid.

Offer: This action indicates that the agent proposes a new bid.

End Negotiation: This action indicates that the agent terminates the entire negotiation, resulting in the lowest possible score for both agents.

If the action was an *Offer*, agent X is subsequently asked to determine its next action and the turn taking goes to the next round. If it is not an *Offer*, the negotiation has finished. The turn taking stops and the final score (utility of the last bid) is determined for each of the agents, as follows:

- The action of agent X is an *Accept*. This action is possible only if the opponent actually did a bid. The last bid of the opponent is taken, and the utility of that bid is determined in the utility spaces of agents A and B .
- The action is returned an *EndNegotiation*. The score of both agents is set to the lowest score.

The parties negotiate over *issues*, and every issue has an associated range of alternatives or *values*. A negotiation outcome consists of a mapping of every issue to a value, and the set Ω of all possible outcomes is called the negotiation *domain*. The domain is common knowledge to the negotiating parties and stays fixed during a single negotiation session. Both parties have certain preferences prescribed by a *preference profile* over Ω . These preferences can be modeled by means of a utility function U that maps a possible outcome $\omega \in \Omega$ to a real-valued number in the range $[0, 1]$. In contrast to the domain, the preference profile of the players is private information.

A bid is a set of chosen values $v_1 \dots v_N$ for each of the N issues (I). Each of these values has been assigned an evaluation value $eval(v_i)$ in the utility space. Each issue has been assigned the normalized weight (w_i , $\sum_{i \in I} w_i = 1$) in the utility space. The utility is the weighted sum of the normalized evaluation values.

The utility function of the bid ($\mathbf{v} = (v_1, \dots, v_N)$) is defined as (1).

$$U(\mathbf{v}) = \sum_{i=1}^N w_i \cdot eval(v_i) \quad (1)$$

A negotiation lasts a predefined time in seconds (*deadline*). The time line is normalized, i.e.: time $t \in [0, 1]$, where $t = 0$ represents the start of the negotiation and $t = 1$ represents the deadline. Apart from a deadline, a scenario may also feature discount factors. Discount factors decrease the utility of the bids under negotiation as time passes. Let d in $[0, 1]$ be the discount factor. Let t in $[0, 1]$ be the current normalized time, as defined by the timeline. We compute the discounted utility U_D^t of an outcome ω from the undiscounted utility function U as follows:

$$U_D^t(\omega) = U(\omega) \cdot d^t \quad (2)$$

At $t = 1$, the original utility is multiplied by the discount factor. Furthermore, if $d = 1$, the utility is not affected by time, and such a scenario is considered to be undiscounted.

3.2 Learning from Past Negotiation Sessions

Recently, automated negotiation agents have had the concept introduced that an agent can save and load information for each preference profile. This means that an agent can learn from previous negotiations, against the same opponent or multiple opponents, to improve its competence when having a specific preference profile. By analyzing the past negotiation sessions, agents can estimate the opponent's utility function based on exchanging bids. For example, the bids an opponent proposes many times in the early stage might be the effective bids for the opponents. The last bid proposed by the opponent might be the lowest utility for agreeing with the bid.

The information an agent can save and load for each preference profile and opponent is as follows: Offered bids, received bids,¹ and exchange sequence of the bids. Therefore, we need to predict or analyze the opponent's utility of bids to utilize the past negotiation sessions.

4 Automated Agent Based on Compromise Strategy

This section shows the compromising strategies [12] based on our proposed strategies.

4.1 Opponent Modeling in Basic Strategy

Our agent estimates the alternatives the opponent will offer in the future based on the opponent's offers. In particular, we estimate them using the values mapping the opponent's bids to our own utility function. The agent works at compromising to the estimated optimal agreement point.

Concretely, our behavior is decided based on the following Eqs. (3), (4).

$$emax(t) = \mu(t) + (1 - \mu(t))d(t) \quad (3)$$

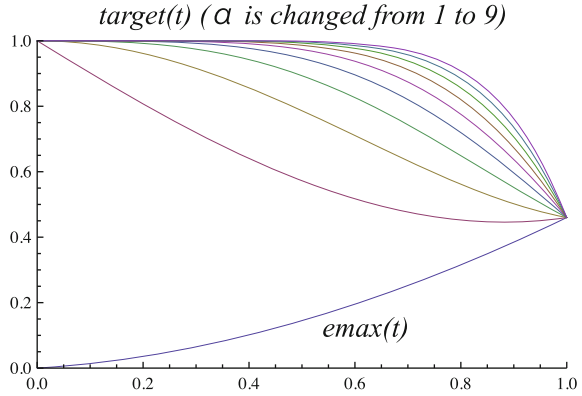
$$target(t) = 1 - (1 - emax(t))t^\alpha \quad (4)$$

$emax(t)$ means the estimated maximum utility of a bid the opponent will propose in the future. $emax(t)$ is calculated by $\mu(t)$ (the mean of the opponent's offers in our utility space), $d(t)$ (the width of the opponent's offers in our utility space) when the timeline is t . $d(t)$ is calculated based on the deviation. We can see how favorable the opponent's offer is based on the deviation ($d(t)$) and the mean ($\mu(t)$).

If we assume that the opponent's offer is generated based on uniform distribution $[\alpha, \alpha + d(t)]$, the deviation is calculated as follows.

¹Bids don't include the utility information.

Fig. 1 $target(t)$ when $emax(t)$ is $\mu(t) = \frac{1}{10}t$ $d(t) = \frac{2}{5}t^2$



$$\sigma^2(t) = \frac{1}{n} \sum_{i=0}^n x_i^2 - \mu^2 = \frac{d^2(t)}{12} \quad (5)$$

Therefore, $d(t)$ is defined as follows.

$$d(t) = \sqrt{12}\sigma(t) \quad (6)$$

We consider the means as the weights for the following reason. When the mean of the opponent's action is located at the center of the domain of the utility, $emax(t)$ is the mean plus half of the width of the opponent's offers. However, it is possible to move only in the high direction when the mean of the utility value is low, and the action can be expanded only in the low direction when the mean is high. Therefore, an accurate estimation is made by introducing the weights.

$target(t)$ is a measure of proposing a bid when time is t , and α is a coefficient for adjusting the speed of compromise. It is effective to search for the opponent's utility information by repeating the proposal to each other as long as time allows. On the other hand, our utility value is required to be as high as possible. Our bids are the higher utility for the opponent at the first stage, and approach asymptotically to $emax(t)$ as the number of negotiation rounds increases.

Figure 1 is an example of $target(t)$ when α is changed from 1 to 9. $emax(t)$ is $\mu(t) = \frac{1}{10}t$, $d(t) = \frac{2}{5}t^2$.

4.2 Proposal and Response Opponent's Bids

First, we show the method of selecting the bids from our utility space. Our agent searches for alternatives whose utility is $target(t)$ by changing the starting points randomly by iteratively deepening the depth-first search method. Next, we show the decision of whether to accept the opponent's offer. Our agent judges whether

to accept it based on $target(t)$ and the mean of the opponent's offers. Equation (7) defines the probability of acceptance.

$$P = \frac{t^5}{5} + (Offer - emax(t)) + (Offer - target(t)) \quad (7)$$

Acceptance probability P is calculated using t , $Offer$, $target(t)$ and the estimated maximum value $emax(t)$. $Offer$ is the utility of the opponent's bid in our utility space.

5 Strategy Adaptation Based on Past Negotiation Sessions

The compromising strategy described in the previous section has following issues:

1. Determination of α adjusting the speed of compromising isn't easy.
2. It doesn't always find the pareto optimal bids in searching bids.

To solve these issues, we propose two strategies using past negotiation sessions.

5.1 Adaptation Strategies Using Past Negotiation Sessions

An opponent's strategy is predictable based on earlier encounters or an experience profile, and can be characterized in terms of some global style, such as the negotiation styles [20, 21], or a known conflict-handling style. One important style is the Thomas-Kilmann Conflict Mode Instrument (TKI) [13, 17]. The TKI is designed to measure a person's behavior in conflict situations. "Conflict situations" are those in which the concerns of two people appear to be incompatible. In this situation, an individual's behavior has two dimensions: (1) assertiveness, the extent to which the person attempts to satisfy his own concerns, and (2) cooperativeness, the extent to which the person attempts to satisfy the other person's concerns. These two basic dimensions of behavior define five different modes for responding to conflict situations: Competing, Accommodating, Avoiding, Collaborating, and Compromising as Fig. 2 shows.

The left side of Table 1 shows the relationships between the condition and cooperativeness, and the right side of Table 1 shows the relationship between the condition and assertiveness. When bid_t (opponent's bid in time t) is higher than μ_h (mean of the bids from past negotiation sessions), our agent regards the opponent as uncooperative. On the other hand, when bid_t is lower than μ_h , our agent regards the opponent as cooperative. In addition, our agent evaluates the assertiveness by comparing between the variance of proposals in the session and that in past negotiation sessions. Usually, assertive agents tend to propose the same bids because they try to push through their proposals by proposing many times. In other words, it is hard for our agent to make win-win agreements when the opponent's bids are disspread. On the other hand,

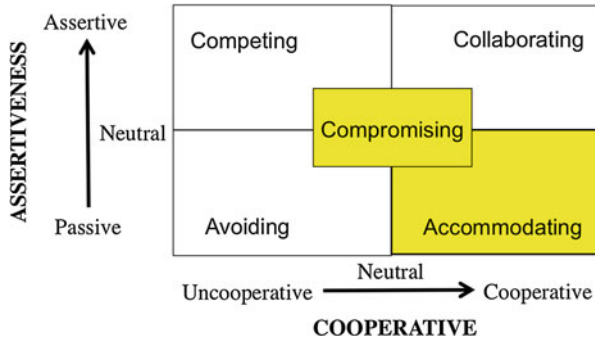


Fig. 2 Overview of Thomas-Kilmann conflict mode instrument (TKI)

Table 1 Estimation of cooperativeness and assertiveness based on past negotiation sessions

| Condition | Cooperativeness | Condition | Assertiveness |
|--------------------|-----------------|----------------------------|---------------|
| $u(bid_t) > \mu_h$ | Uncooperative | $\sigma^2(t) > \sigma_h^2$ | Passive |
| $u(bid_t) = \mu_h$ | Neutral | $\sigma^2(t) = \sigma_h^2$ | Neutral |
| $u(bid_t) < \mu_h$ | Cooperative | $\sigma^2(t) < \sigma_h^2$ | Assertive |

passive agents tend to propose various bids because they change their proposals by searching for win-win agreements. In other words, our agent can make an agreement when the opponent's bids are spread. Considering the above theory, our agent tries to compromise more and more when the opponent is cooperative and passive, which means the opponent is "accommodating" or "compromising" (yellow box in Fig. 2) in the TKI. For judging the opponent's TKI, we employ the past negotiation sessions.

Figure 3 shows the concept of adjusting the speed of compromising in this paper. As Eq. (4) in Sect. 4.1 shows, the speed of compromising is decided by α in $target(t)$. α is set as a higher value at the first stage, and α is decreased when the opponent is "accommodating" or "compromising." By introducing this adjustment algorithm, our agent can adjust its strategy from hardheaded to cooperative more and more when the opponent tries to make agreement. When there is a discount factor, our agent can make an agreement in the early stage by employing the adjustment of α , despite that the existing compromising strategy makes an agreement just before the finish. In addition, our agent can prevent poor compromising because it considers the opponent's strategy and situation.

The detailed algorithm of adapting the agent's strategies based on past negotiation sessions is as follows:

1. Our agent sets α in $target(t)$ to the highest value.
2. It calculates the mean (μ_h) and variance (σ_h^2) of the opponent's bids from past negotiation sessions in appropriate domains.
3. It calculates the utility of offered bid in time t ($u(bid_t)$) and the variance of offered bids from 0 to t ($\sigma^2(t)$).

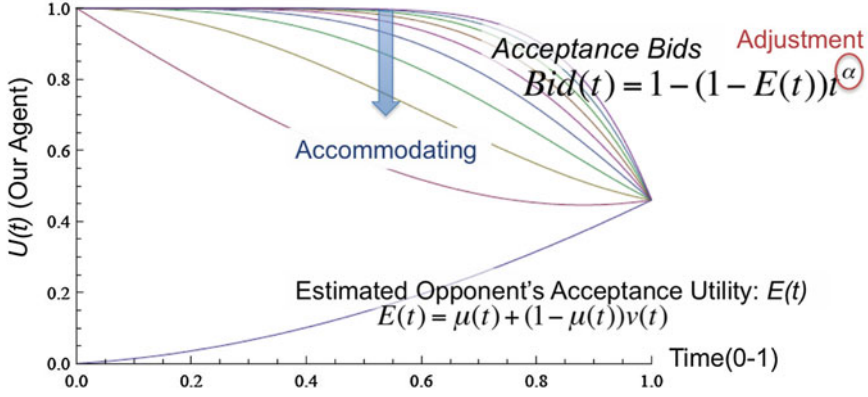


Fig. 3 Adjustment of speed of compromising

4. It compares between μ_h and $u(bid_t)$ to judge the cooperativeness.
5. It compares between σ_h^2 and $\sigma^2(t)$ to judge the assertiveness.
6. It updates the α in $target(t)$ based on the following equation when the opponent is “accommodating” or “compromising”:

$$\alpha' = \alpha - \varepsilon \quad (8)$$

(α' is a renewed coefficient for adjusting the speed of compromise, ε is a constant for adjusting the α .)

5.2 Searching for Pareto Optimal Bids

The proposed agent can search for pareto optimal bids based on the similarity between bids. The opponents don't reveal their preferences to each other in the negotiation; therefore, it isn't easy for agents to search for the pareto optimal bids. In this paper, the agent tries to find the bids that are similar to the opponent's first bid because the first bid has high possibility of being the best bid for the opponent.

In this paper, our agent tries to find the most similar bids using the following equation. \mathbf{v}_0 means the opponent's bid proposed the first time, and \mathbf{v}_x means the target bid for evaluating the similarity. The similarity between \mathbf{v}_0 and \mathbf{v}_x ($sim(\mathbf{v}_0, \mathbf{v}_x)$) is defined as follows:

$$sim(\mathbf{v}_0, \mathbf{v}_x) = \sum_{i=1}^m w_i \cdot bool(v_0, v_i) \quad (9)$$

($bool(v_0, v_i)$: if ($v_0 == v_i$) then return 1 else return 0)

Our agent searches for the bids in which the utility is the same as $target(t)$ and $sim(\mathbf{v}_0, \mathbf{v}_x)$ is highest using the repeated depth-first search algorithm.

6 Experimental Analysis

The performance of our proposed agent is evaluated with GENIUS (General Environment for Negotiation with Intelligent multipurpose Usage Simulation [15]), which is also used as a competition platform for ANAC.

6.1 Qualifying Round Results of ANAC-2013

Nineteen agents were submitted to the competition. The 11 domains were selected from archives submitted by the participants of ANAC-2013. For each pair of agents, under each utility function, we ran a total of 20 negotiations (including the exchange of preference profiles). In other words, 75,240 sessions are run in the qualifying round. The maximum negotiation time of each negotiation session is set to 3 min and normalized into the range of $[0, 1]$. Table 2 shows mean scores over all the scores achieved by each agent (against all opponents and using all utility functions) and variances.

Note that these means are taken over all negotiations, excluding those in which both agents use the same strategy (i.e. excluding self-play). Therefore, the mean score $U_\Omega(p)$ of agent p in scenario Ω is given formally by:

$$U_\Omega(p) = \frac{\sum_{p' \in P, p \neq p'} U_\Omega(p, p')}{(|P| - 1)} \quad (10)$$

where P is the set of players and $U_\Omega(p, p')$ is the utility achieved by player p against player p' in scenario Ω . For every domain, due to the normalization of the scores, the lowest possible score is 0 and the highest is 1. The fact that the maximum and minimum scores are not always achieved can be explained by the non-deterministic behavior of the agents: the top-ranking agent on one domain does not always obtain the maximum score on every trial.

As Table 2 shows, our agent has won by a big margin in the qualifying round of ANAC-2013. Considering the variance among the domains, our agent had advantages compared with other agents. Some reasons for this are as follows. First, we try to improve the speed of making agreements by adjusting $emax(t)$. In addition, our agent tries to compromise positively when the opponent is cooperative. Agents couldn't learn from the past negotiation sessions in the past ANAC; therefore, they tried to find effective agreements by eliciting the opponent's utility in the negotiation session. In other words, agents won the utility decreased by the discount factor

Table 2 Results of every combination among ANAC-2013 agents

| | Agent | Rank | Mean | Variance |
|----|---------------------|----------|--------------|----------------|
| 1 | Our Agent | 1 | 0.562 | 0.00019 |
| 2 | Agent Slinkhard | 2–3 | 0.522 | 0.00132 |
| 3 | TMFAgent | 2–4 | 0.516 | 0.00163 |
| 4 | MetaAgent | 3–4 | 0.495 | 0.00252 |
| 5 | GAgent | 5–8 | 0.457 | 0.00241 |
| 6 | InoxAgent | 5–8 | 0.455 | 0.00235 |
| 7 | SlavaAgent | 5–11 | 0.447 | 0.00018 |
| 8 | VAStockMarketAgent | 5–11 | 0.446 | 0.0052 |
| 9 | RoOAgent | 7–11 | 0.432 | 0.00313 |
| 10 | AgentTalex | 7–11 | 0.431 | 0.00285 |
| 11 | AgentMRK2 | 7–11 | 0.43 | 0.00344 |
| 12 | Elizabeth | 12–14 | 0.387 | 0.00443 |
| 13 | ReuthLiron | 12–15 | 0.374 | 0.00416 |
| 14 | BOAconstrictorAgent | 12–15 | 0.373 | 0.00141 |
| 15 | Pelican | 13–18 | 0.359 | 0.00434 |
| 16 | Oriel_Einat_Agent | 15–18 | 0.35 | 0.00534 |
| 17 | MasterQiao | 15–18 | 0.345 | 0.00214 |
| 18 | Eagent | 15–18 | 0.338 | 0.00707 |
| 19 | ClearAgent | 19 | 0.315 | 0.00109 |

because they needed to continue many rounds to get enough of the opponent’s utility information. However, our agent tries to make agreements in the early stage using the past negotiation sessions when the opponent looks cooperative. Second, our agent could propose pareto optimal bids many times. If agents could offer the pareto optimal bids, the offers are effective and easy for making win-win agreements. Therefore, our agent could find better agreements by the effective search technique.

6.2 Detailed Experimental Analysis

We compare the negotiation efficiency of our proposed agent with eight state-of-the-art negotiation agents that entered the final round of ANAC-2013: GAgent, MetaAgent, SlavaAgent, TMFAgent, which are implemented by negotiation experts from different research groups.² In addition, we added the AgentK, which strategy is the

²All the agents and the domains that participated in the final round of ANAC-2013 are available in GENIUS 4.2.

Table 3 Mean utility of agreement of each agent against all opponents in different domains

| Agent | Acquisition | | HouseKeeping | |
|------------------|----------------|----------------|----------------|----------------|
| | Mean | Variance | Mean | Variance |
| Our Agent | 0.85198 | 0.00681 | 0.47304 | 0.08937 |
| GAgent | 0.90448 | 0.04179 | 0.44034 | 0.061785 |
| Meta-Agent | 0.91750 | 0.05772 | 0.54888 | 0.04447 |
| Slava-Agent | 0.94569 | 0.05568 | 0.41771 | 0.051393 |
| TMFAgent | 0.92361 | 0.00639 | 0.43353 | 0.08918 |
| AgentK | 0.87369 | 0.01248 | 0.38956 | 0.047646 |

Table 4 Mean time of agreement of each agent against all opponents in different domains

| Agent | Acquisition | | HouseKeeping | |
|------------------|----------------|----------------|----------------|----------------|
| | Mean | Variance | Mean | Variance |
| Our Agent | 0.33296 | 0.04595 | 0.22151 | 0.17172 |
| GAgent | 0.64804 | 0.04986 | 0.50416 | 0.078815 |
| Meta-Agent | 0.84832 | 0.06358 | 0.90832 | 0.073640 |
| Slava-Agent | 0.62320 | 0.04830 | 0.68524 | 0.08104 |
| TMFAgent | 0.98483 | 0.00425 | 0.54105 | 0.076666 |
| AgentK | 0.68889 | 0.03910 | 0.62783 | 0.073640 |

Table 5 Mean distance to Pareto frontier for each agent against all opponents in different domains

| Agent | Acquisition | HouseKeeping |
|------------------|------------------|----------------|
| | Mean | Mean |
| Our Agent | 0.0021133 | 0.13090 |
| GAgent | 0.093006 | 0.19863 |
| Meta-Agent | 0.035986 | 0.03199 |
| Slava-Agent | 0.12254 | 0.44346 |
| TMFAgent | 0.00013182 | 0.0 |
| AgentK | 0.0054941 | 0.33268 |

basic compromise strategy [12].³ For each pair of agents, under each utility function, we ran a total of 20 negotiations (including the exchange of preference profiles).

The negotiation domains can be classified based on the characteristic of weak and strong opposition [1]. Domains with strong opposition mean that the agents have strongly opposite interests over the negotiation outcomes and the gain of one agent must come at the loss of the other agent. Domains with weak opposition refer to those domains in which it is possible for the agents to reach a win-win agreement. In this experiment, we consider two different types of negotiation domains: HouseKeeping

³For showing the effectiveness of our improvements, AgentK was included in the experiments.

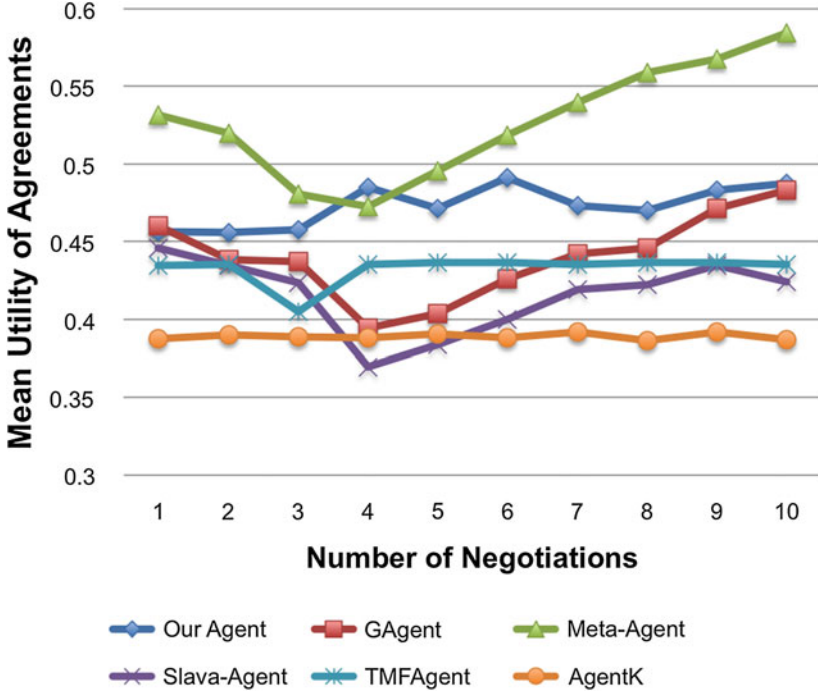


Fig. 4 Changes of mean utility of each agent against all opponents in each round (HouseKeeping)

domain with strong opposition and Acquisition domain with weak opposition. The total number of possible agreements in this domain is 384, and the discount factor and the reservation value of the HouseKeeping domain are set to 0.25 and 0, respectively. Another domain, the Acquisition domain, is much less competitive compared with HouseKeeping. The total number of possible agreements in this domain is 480, and the discount factor and the reservation value are set to 1.0 and 0, respectively. These domains are almost same size; however the opposition and the discount factor between them are totally different.

Table 3 shows the mean normalized scores of agreements for each agent in each domain. Table 4 shows the mean time of agreements for each agent in each domain. Table 5 shows the mean distance to the Pareto frontier for each agent in each domain. The mean distance to the Pareto frontier is defined formally as:

$$\text{meanParetoDistance}(\Omega) = \sum_{\omega \in \text{Offers}} \frac{\min_{\omega_p \in \Omega_p} \text{dist}(\omega, \omega_p)}{|\text{Offers}|}$$

where $\Omega_p \subset \Omega$ is the set of Pareto efficient possible outcomes, and *Offers* is the set of all bids offered by the agent. The ‘dist’ function gives the Euclidean distance between two points in the outcome space, defined formally as:

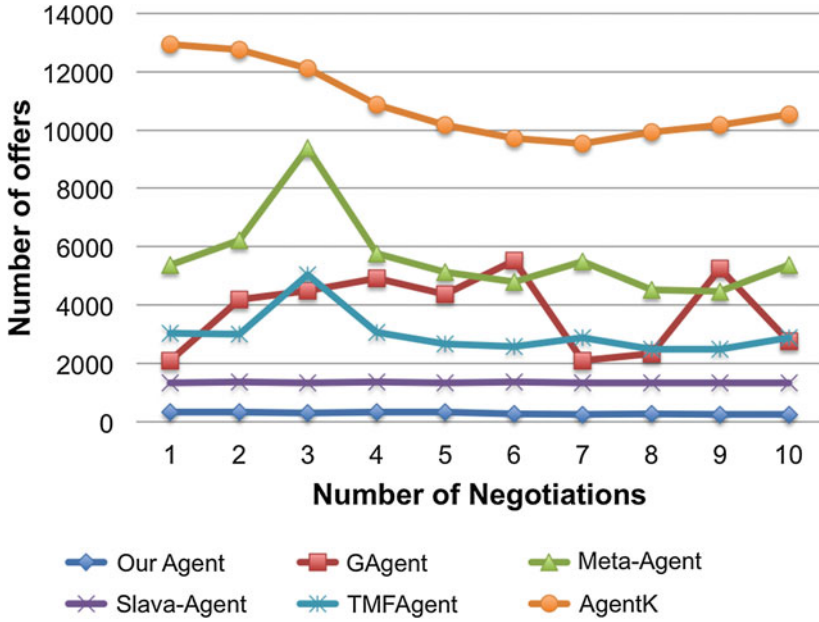


Fig. 5 Changes of mean number of offers of each agent against all opponents in each round (HouseKeeping)

$$\text{dist}(\omega_1, \omega_2) = \sqrt{(U_1(\omega_1) - U_1(\omega_2)) * (U_2(\omega_1) - U_2(\omega_2))}$$

where $U_1(\cdot)$ and $U_2(\cdot)$ give the utilities to players 1 and 2, respectively.

As Table 3 shows, our agent outperforms others expected for Meta-agent with a small variance in the HouseKeeping domain. The main reasons are shown in the results in Tables 4 and 5. Our agent tries to improve the speed of making agreements by adjusting α in $emax(t)$, and compromises positively when the opponent is cooperative. The results of mean time of agreements outperformed compared with other agents in Table 4, definitely. In addition, our agent shows better results than some other agents in Table 5. In other words, these results show the effectiveness of a search method for pareto optimal bids. On the other hand, our agent lost to others in the Acquisition domain because this domain don't have a time-discounting factor.

Figures 4 and 5 show the changes of the mean utility and time of agreements in the HouseKeeping domain. Usually, the results of utility show an inverted parabolic curve in most agents. However, our agent has almost the same utility when the number of negotiations increases. The reasons for this are shown in Fig. 5. Because of the discount factor, the mean utility is influenced a great deal by the mean time of agreement. In fact, the results of mean time show inverted parabolic curves in the all agents except for our agent. The reason for this, shown in Fig. 5, is that we adjust the effective strategy based on TKI. In other words, we can select an effective strategy when the opponent changes its strategies as the negotiation rounds proceed.

7 Conclusion

This paper focused on bilateral multi-issue closed negotiation, which is an important class of real-life negotiations. This paper proposed a novel agent that estimates the alternatives the opponent offers based on past negotiation sessions. In addition, our agent could adjust the speed of compromising using the past negotiation sessions. We demonstrated that the proposed method results in good outcomes and greater search technique for the pareto frontier. Additionally, we demonstrated the change of the utility in multi-times negotiation for analyzing the learning strategies.

In our possible future works, we will prove the amount of past negotiation sessions for judging the opponent's TKI mode. In learning technology (especially real-time learning), cold start problems are important. For proposing and analyzing this issue, we will demonstrate experimentally or prove in theory the amount of past negotiation sessions. In addition, we will prove the timing of changing the strategy in theory. By getting the payoff table every time, the optimal timing of adjusting the agent's strategy can be calculated.

References

1. Baarslag, T., Fujita, K., Gerding, E., Hindriks, K., Ito, T., Jennings, N.R., Jonker, C., Kraus, S., Lin, R., Robu, V., Williams, C.: Evaluating practical negotiating agents: results and analysis of the 2011 international competition. *Artif. Intell. J. (AIJ)* **198**, 73–103 (2013)
2. Baarslag, T., Hindriks, K.: Accepting optimally in automated negotiation with incomplete information. In: *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS2013)*, pp. 715–722 (2013)
3. Chen, S., Weiss, G.: An efficient and adaptive approach to negotiation in complex environments. In: *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI-2012)*, vol. 242, pp. 228–233 (2012)
4. Chen, S., Weiss, G.: An efficient automated negotiation strategy for complex environments. *Eng. Appl. Artif. Intell.* (2013). doi:[10.1016/j.engappai.2013.08.012](https://doi.org/10.1016/j.engappai.2013.08.012)
5. Faratin, P., Sierra, C., Jennings, N.R.: Using similarity criteria to make issue trade-offs in automated negotiations. *Artif. Intell.* **142**, 205–237 (2002)
6. Fatima, S.S., Wooldridge, M., Jennings, N.R.: Multi-issue negotiation under time constraints. In: *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2002)*, pp. 143–150. New York, NY, USA (2002)
7. Fatima, S.S., Wooldridge, M., Jennings, N.R.: Multi-issue negotiation with deadlines. *J. Artif. Intell. Res. (JAIR)* **27**, 381–417 (2006)
8. Gal, K., Ito, T., Jonker, C., Kraus, S., Hindriks, K., Lin, R., Baarslag, T.: The forth international automated negotiating agents competition (anac2013). <http://www.itolab.nitech.ac.jp/ANAC2013/> (2013)
9. Hao, J., Leung, H.F.: Abines: An adaptive bilateral negotiating strategy over multiple items. In: *2012 IEEE/WIC/ACM International Conferences on Intelligent Agent Technology (IAT-2012)*, vol. 2, pp. 95–102 (2012). doi:[10.1109/WI-IAT.2012.72](https://doi.org/10.1109/WI-IAT.2012.72)
10. Ito, T., Jonker, C., Kraus, S., Hindriks, K., Fujita, K., Lin, R., Baarslag, T.: The second international automated negotiating agents competition (anac2011). <http://www.anac2011.com/> (2011)

11. Jonker, C., Kraus, S., Hindriks, K., Lin, R., Tykhonov, D., Baarslag, T.: The first automated negotiating agents competition (anac2010). <http://mmi.tudelft.nl/negotiation/tournament> (2010)
12. Kawaguchi, S., Fujita, K., Ito, T.: Compromising strategy based on estimated maximum utility for automated negotiation agents competition (anac-10). In: 24th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE-2011), pp. 501–510 (2011)
13. Kilmann, R.H., Thomas, K.W.: Developing a forced-choice measure of conflict-handling behavior: the mode instrument. *Eng. Appl. Artif. Intell.* **37**(2), 309–325 (1977)
14. Kraus, S.: *Strategic Negotiation in Multiagent Environments*. MIT Press (2001)
15. Lin, R., Kraus, S., Baarslag, T., Tykhonov, D., Hindriks, K., Jonker, C.M.: Genius: an integrated environment for supporting the design of generic automated negotiators. *Comput. Intell.* (2012). doi:[10.1111/j.1467-8640.2012.00463.x](https://doi.org/10.1111/j.1467-8640.2012.00463.x)
16. Osborne, M.J., Rubinstein, A.: *Bargaining and Markets (Economic Theory, Econometrics, and Mathematical Economics)*. Academic Press (1990)
17. Pruitt, D.G.: Trends in the scientific study of negotiation and mediation. *Negot. J.* **2**(3), 237–244 (1986). doi:[10.1111/j.1571-9979.1986.tb00361.x](https://doi.org/10.1111/j.1571-9979.1986.tb00361.x)
18. Rubinstein, A.: Perfect equilibrium in a bargaining model. *Econometrica* **50**(1), 97–109 (1982)
19. Rubinstein, A.: A bargaining model with incomplete information about time preferences. *Econometrica* **53**(5), 1151–1172 (1985)
20. Shell, G.R.: *Bargaining for Advantage: Negotiation Strategies for Reasonable People*. Penguin Books (2006)
21. Thompson, L.: *Mind and Heart of the Negotiator*, 2nd edn. Prentice Hall Press (2000)
22. Williams, C.R., Robu, V., Gerding, E., Jennings, N.R., Ito, T., Jonker, C., Kraus, S., Hindriks, K., Lin, R., Baarslag, T.: The third automated negotiating agents competition (anac2012). <http://anac2012.ecs.soton.ac.uk> (2012)
23. Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R.: Using gaussian processes to optimise concession in complex negotiations against unknown opponents. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011), pp. 432–438 (2011)

Recent Advances in Agent-based Complex Automated
Negotiation

Fukuta, N.; Ito, T.; Zhang, M.; Fujita, K.; Robu, V. (Eds.)

2016, XIII, 293 p. 86 illus., 79 illus. in color., Hardcover

ISBN: 978-3-319-30305-5