TKI Adaptation Strategy for Complex Multi-times Bilateral Negotiations

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

***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. Recently, the attention of this study has focused on the nonlinear utility functions. In nonlinear utility functions, most of the negotiation strategies for linear utility functions can't adopt to the scenarios of nonlinear utility functions. 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 Mode and search for the pareto frontier using past negotiation sessions. In the experiments, we demonstrate that the proposed agent has better outcomes and greater search technique than existing agents.**

***Keywords*-Multi-issue Negotiation; Multi-agent System; Agreement Technology**

I. 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. [1]). 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 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 [2]. 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 [3]. Also, some effective automated negotiating agents have been proposed through the competitions [4].

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. Another key point in achieving automated negotiation in real life is the non-linearity of the utility functions. Many real-world negotiation problems assume the multiple nonlinear utility function. When an automated negotiation strategy covers the linear function effectively, it is not always possible or desirable in the nonlinear situations [5].

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) [6]. 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. In addition, our agent has an effective search strategy for finding the pareto optimal bids. The main idea of this strategy was proposed by our group [7], however, which focused on the linear utility functions only. In addition, the another paper tries to adopt this strategy to the nonlinear situations [8]. However, the search method of this paper was poor, therefore, the proposed method can't find the effective bids in the large size domain. This paper focuses on the nonlinear utility function and employs the effective search technique, which is close to the negotiation in the real life.

In the experiments, we demonstrate that the proposed agent outperforms the other agents that participated in the final round of ANAC-2014. 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

978-1-4673-9469-7/15 $31.00 © 2015 IEEE
DOI 10.1109/SOCA.2015.30

in the earlier steps in the nonlinear domains.

The remainder of the paper is organized as follows. First, we describe related works. Next, 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 tournaments among top-4 finalist in ANAC-2014 and some experimental analysis. Finally, we present our conclusions.

II. Negotiation Environments

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 [9], in which the negotiating parties exchange offers in turns. The alternating-offers protocol conforms with our criterion to have simple rules.

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.

An agent's utility function, in the formulation, is described in terms of constraints [10]. There are l constraints, $c_k \in C$. Each constraint represents a region in the contract space with one or more dimensions and an associated utility value. In addition, c_k has value $v_a(c_k, \vec{s})$ if and only if it is satisfied by contract $\vec{s}$. Every agent has its own, typically unique, set of constraints. An agent's utility for contract $\vec{s}$ is defined as the weighted sum of the utility for all the constraints it satisfies, i.e., as $u_a(\vec{s}) = \sum_{c_k \in C, \vec{s} \in x(c_k)} v_a(c_k, \vec{s})$, where $x(c_k)$ is a set of possible contracts (solutions) of c_k. This expression produces a "bumpy" nonlinear utility function with high points where many constraints are satisfied and lower regions where few or no constraints are satisfied. This represents a crucial departure from previous efforts on multi-issue negotiation, where contract utility is calculated as the weighted sum of the utilities for individual issues, producing utility functions shaped like flat hyperplanes with a single optimum.

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: $U_D^t(\omega) = U(\omega) \cdot d^t$ 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.

In addition, 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.

III. Compromise Strategy for Negotiaitons

This section shows the compromising strategies [4], which is based on our proposed strategies.

A. 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 equations (1), (2).

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

$$target(t) = 1 - (1 - emax(t))t^{\alpha} \quad (2)$$

$emax(t)$ 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 deviation of the opponent's offers in our utility space. In other words, it means the width of the opponent's offers in our utility space) when the timeline is t.

$target(t)$ 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

[1] Bids don't include the utility information.

opponent at the first stage, and approach asymptotically to $emax(t)$ as the number of negotiation rounds increases.

B. 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 (3) defines the probability of acceptance.

$$P = \frac{t^5}{5} + (Offer - emax(t)) + (Offer - target(t)) \quad (3)$$

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.

IV. Strategy Adaptation and Efficient Search Technique based on Past Negotiation Sessions

A. Strategy Adaptation using TKI

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.

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 styless, or a known conflict-handling style. One important style is the Thomas-Kilmann Conflict Mode Instrument (TKI) [6], [11]. 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.

The left side of Table I shows the relationships between the condition and cooperativeness, and the right side of Table I 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

Table I
Estimation of Cooperativeness and Assertiveness based on Past Negotiation Sessions

Condition	Cooperativeness
$u(bid_t) > \mu_h$	Uncooperative
$u(bid_t) = \mu_h$	Neutral
$u(bid_t) < \mu_h$	Cooperative

Condition	Assertiveness
$\sigma^2(t) > \sigma_h^2$	Passive
$\sigma^2(t) = \sigma_h^2$	Neutral
$\sigma^2(t) < \sigma_h^2$	Assertive

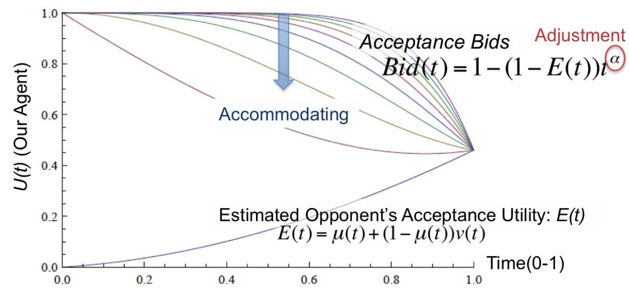

Figure 1. Adjustment of Speed of Compromising

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 dispread. On the other hand, 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" in the TKI. For judging the opponent's TKI, we employ the past negotiation sessions.

Figure 1 shows the concept of adjusting the speed of compromising in this paper. As equation (2) in the previous section 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 ($\sigma^2(t)$).
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":

$$\alpha' = \alpha - \epsilon \tag{4}$$

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

B. 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. $\vec{v_0}$ means the opponent's bid proposed the first time, and $\vec{v_x}$ means the target bid for evaluating the similarity. The similarity between $\vec{v}_0$ and $\vec{v}_x(sim(\vec{v}_0, \vec{v}_x))$ is defined as follows:

$$sim(\vec{v}_0, \vec{v}_x) = \sum_{i=1}^{m} w_i \cdot bool(v_0, v_i) \tag{5}$$

($bool(v_0, v_i)$: *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 $sim(\vec{v}_0, \vec{v}_x)$ is highest using the following methods: Exhaustive Search with limitations and Simulated Annealing (SA) [12]. Exhaustive Search with limitations searches the all possible bids, exhaustively. However, exhaustive search can't find the bids in the large-sized domains because it becomes computationally intractable as the search space grows. In the small-size nonlinear utility function, Random Sampling can find the optimal solution. In the large-size nonlinear utility function, SA is more effective to search the optimal bids.

V. Experimental Analysis

The performance of our proposed agent is evaluated with GENIUS (General Environment for Negotiation with Intelligent multipurpose Usage Simulation [13]), which is also used as a competition platform for ANAC. We evaluated our agent by comparing with four state-of-the-art agents submitted in ANAC-2014 (AgentM, E2Agent, Gangster, and Whale Agent[2]). These four agents are the top-4 in ANAC-2014 social utility categories, in other words, they are effective and state-of-the-art agents to get the high social utility in the nonlinear utility functions. They are implemented by negotiation experts from different research groups.

Table II
INDIVIDUAL UTILITY OF EACH AGENT AGAINST OUR AGENT IN DIFFERENT NEGOTIATION DOMAINS

Agent	10^{10} domain	10^{30} domain	10^{50} domain
OurAgent (Exhaustive)	**0.663584**	**0.706638**	**0.692816**
AgentM	0.557693	0.709559	0.704370
OurAgent (Exhaustive)	**0.476070**	**0.151085**	**0.301058**
E2Agent	0.400251	0.133538	0.369168
OurAgent (Exhaustive)	**0.649159**	**0.512566**	**0.522964**
WhaleAgent	0.529163	0.395057	0.467015
OurAgent (Exhaustive)	**0.391714**	**0.225447**	**0.241957**
Gangster	0.365660	0.207486	0.259846

Table III
INDIVIDUAL UTILITY OF EACH AGENT AGAINST OUR AGENT IN DIFFERENT NEGOTIATION DOMAINS

Agent	10^{10} domain	10^{30} domain	10^{50} domain
OurAgent (SA)	**0.846037**	**0.638043**	**0.537344**
AgentM	0.853416	0.632626	0.535427
OurAgent (SA)	**0.711508**	**0.405362**	**0.385359**
E2Agent	0.741149	0.358858	0.363256
OurAgent (SA)	**0.664950**	**0.377130**	**0.374023**
WhaleAgent	0.710123	0.371275	0.369920
OurAgent (SA)	**0.684588**	**0.648502**	**0.604359**
Gangster	0.714015	0.500792	0.486714

The three domains were selected from archives of ANAC-2014. The size of the domains are $10^{10}, 10^{30}, and\ 10^{50}$. Each constraint in these domains are related to 1 to 5 issues. The horizontal axis means the agent A's utility and vertical axis means the agent B's utility in each figure. The three scenarios contained broadly similar characteristics such as the shapes of the pareto frontier and so on. In all domains, the discount factor and the reservation value are set to 0.5 and 0, respectively. For each pair of agents, under each utility function (including the exchange of preference profiles), repeating 10 times under the same opponents and utility functions. The maximum negotiation time of each negotiation session is set to 3 minutes and normalized into the range of $[0, 1]$.

In "Our Agent(SA)," the SA initial temperature was 20.0 and decreases linearly to 0 over the course of 200 iterations. The initial contract for each SA run is randomly selected.

Table II shows mean scores over all the scores achieved

[2]All the agents and the domains that participated in the final round of ANAC-2014 will be available in the newest GENIUS.

by each agent against Our Agent (Exhaustive), and Table III shows mean scores over all the scores achieved by each agent against Our Agent (SA). As Table II shows, our agent(Exhaustive) has won by a big margin in the 10^{10} domain and 10^{30} domain. 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 because they needed to continue many rounds to get enough of the opponent's utility information. On the other hand, 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(Exhaustive) could find better agreements by the effective search technique. However, our agent(Exhaustive) couldn't win in the large size domain such as the 10^{50} domain because our agent(Exhaustive) has limitation of searching, and needs a lot of time to finish searching in the large-size domains.

As table III shows, our agent(SA) has won by a big margin in the 10^{30} domain and 10^{50} domain. This is because that our agent could find the optimal solution rapidly in the large-size domains with the effectiveness of our agent's strategies. However, our agent(SA) couldn't win in the small size domain such as the 10^{10} domain because our agent(SA) misses some important bids by using the search technique with the random nature in the small-size domain. In addition, the local-optimal in the agent's utility space are appeared because the utility function with many constraints in the small-size domains tends to be bumpy.

VI. Conclusion

This paper focused on bilateral multi-issue closed "nonlinear" 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 with SA have good outcomes in the large-size domains. 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] S. Kraus, *Strategic Negotiation in Multiagent Environments*. Mit Press, 2001.

[2] K. Gal, T. Ito, C. Jonker, S. Kraus, K. Hindriks, R. Lin, and T. Baarslag, "The forth international automated negotiating agents competition (anac2013)," http://www.itolab.nitech.ac.jp/ANAC2013/, 2013.

[3] T. Baarslag, K. Fujita, E. Gerding, K. Hindriks, T. Ito, N. R. Jennings, C. Jonker, S. Kraus, R. Lin, V. Robu, and C. Williams, "Evaluating practical negotiating agents: Results and analysis of the 2011 international competition," *Artificial Intelligence Journal (AIJ)*, vol. 198, pp. 73–103, 2013.

[4] S. Kawaguchi, K. Fujita, and T. Ito, "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)*, 2011, pp. 501–510.

[5] M. Klein, P. Faratin, H. Sayama, and Y. Bar-Yam, "Negotiating complex contracts," *Group Decision and Negotiation*, vol. 12, no. 2, pp. 58–73, 2003.

[6] R. H. Kilmann and K. W. Thomas, "Developing a forced-choice measure of conflict-handling behavior: The mode instrument," *Engineering Applications of Artificial Intelligence*, vol. 37, no. 2, pp. 309–325, 1977.

[7] K. Fujita, "Automated strategy adaptation for multi-times bilateral closed negotiations," in *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014)*, 2014, pp. 1509–1510.

[8] ——, "Compromising adjustment based on conflict mode for multi-times bilateral closed nonlinear negotiations," in *17th International Conference of Principles and Practice of Multi-Agent Systems (PRIMA2014)*, 2014, pp. 439–454.

[9] A. Rubinstein, "Perfect equilibrium in a bargaining model," *Econometrica*, vol. 50, no. 1, pp. 97–109, 1982.

[10] T. Ito, H. Hattori, and M. Klein, "Multi-issue negotiation protocol for agents : Exploring nonlinear utility spaces," in *Proc. of the 20th International Joint Conference on Artificial Intelligence (IJCAI-2007)*, 2007, pp. 1347–1352.

[11] D. G. Pruitt, "Trends in the scientific study of negotiation and mediation," *Negotiation Journal*, vol. 2, no. 3, pp. 237–244, 1986.

[12] S. J. Russell and P. Norvig, *Artificial Intelligence : A Modern Approach*. Prentice Hall, 2002.

[13] R. Lin, S. Kraus, T. Baarslag, D. Tykhonov, K. Hindriks, and C. M. Jonker, "Genius: An integrated environment for supporting the design of generic automated negotiators," *Computational Intelligence*, 2012.