

Meta-Strategy for Multi-Time Negotiation: A Multi-Armed Bandit Approach

Extended Abstract

Ryohei Kawata and Katsuhide Fujita

Tokyo University of Agriculture and Technology, Tokyo, Japan
kawata@katfuji.lab.tuat.ac.jp, katfuji@cc.tuat.ac.jp

ABSTRACT

Multi-time negotiation, which repeats negotiations many times under the same conditions, is an important class of automated negotiation. We propose a meta-strategy that selects an agent's individual negotiation strategy for multi-time negotiation. We model the meta-strategy as a multi-armed bandit problem that regards an individual negotiation strategy as a slot machine and utility of the agent as a reward. Our meta-strategy takes an individual negotiation strategy according to the opponent's strategy, its own profile, and the opponent's profile. The experimental results demonstrate the effectiveness of our meta-strategy under various negotiation conditions.

KEYWORDS

automated negotiation; meta-strategy; multi-time negotiation; multi-armed bandit problem

ACM Reference Format:

Ryohei Kawata and Katsuhide Fujita. 2019. Meta-Strategy for Multi-Time Negotiation: A Multi-Armed Bandit Approach. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, Montreal, Canada, May 13–17, 2019, IFAAMAS, 3 pages.

1 INTRODUCTION

Automated negotiation resolves conflicts among agents and enables them to cooperate with each other [8, 11, 12]. Multi-time negotiation is an important class of automated negotiation. Multi-time negotiation, which repeats negotiations many times under the same conditions, is one of the remarkable topics in the academic automated negotiations. The setup of multi-time negotiation problem is based on real-life negotiations such as supply chain management and service level agreement. The automated negotiation competition focuses on multi-time negotiation in the recent competition [4, 10].

The performance of a negotiating agent depends strongly on the negotiation situation, which includes factors such as the opponent, domain, and profiles [6, 7, 9]. In other words, no single negotiation strategy clearly outperforms all others in every situation. Therefore, agents should select a suitable strategy to reach a beneficial agreement. In multi-time negotiation, agents can utilize their own negotiation history to select an effective strategy.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada. © 2019 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

2 META-STRATEGY

In this paper, we defined “meta-strategy” as overarching strategy that determines the individual negotiation strategy to use in a given negotiation situation. In addition, an individual negotiation strategy is defined as the elements to compose of the meta-strategy. Our meta-strategy includes some individual negotiation strategies and selects one from them as an agent's negotiation strategy at the beginning of the negotiation. We model the selection of individual negotiation strategy as a multi-armed bandit problem [14] by considering an individual negotiation strategy as a slot machine and its own utility as a reward. We adopt the ϵ -greedy algorithm and the upper confidence bound (UCB) algorithm [1] to the selection. The ϵ -greedy algorithm takes an exploration with a probability of ϵ , and takes an exploitation with a probability of $1 - \epsilon$. The exploration is to select a strategy at random. The exploitation is to select the strategy with the highest average utility in the past negotiations. The UCB algorithm selects the negotiation strategy with the highest UCB score at every selection. Let S represent the set of negotiation strategies, N represent the total number of trials, and N_s is the number of trials for each negotiation strategy $s \in S$. The UCB score of negotiation strategy $s \in S$ is $UCB(s) = \hat{\mu}_s + c\sqrt{\frac{\ln N}{N_s}}$, where $\hat{\mu}$ is the average utility of s in the past trials and c is a parameter that controls the frequency of exploration ($c > 0$).

Ilany and Gal proposed the strategy selection method based on supervised learning using structural features of the negotiation domain [9]. Further, they presented two selection methods: the selection method based on a bandit approach (pure-MAB) and the method that combines pure-MAB and the supervised learning approach (prior-MAB). Although prior-MAB outperforms other single negotiation strategy agents, it needs different types of training data for each class of negotiating problem. In addition, pure-MAB and prior-MAB calculate the UCB score for each its own profile. In other words, they do not consider the opponent's strategy and profile. Our approach calculates the UCB score for each combination of its own profile and the opponent's strategy and profile.

Our agents select a strategy to use in a negotiation from the existing strategies, namely, *Atlas3*, *CaduceusDC16*, *kawaii*, *ParsCat*, *Rubick*, and *YXAgent*. These strategies are outstanding strategies in the individual utility category of ANAC2015, ANAC2016 and ANAC2017 [2, 3, 5]. Each strategy has different characteristics and outperforms the others in several situations.

3 EXPERIMENTS

We conducted experiments to evaluate our meta-strategy. The negotiation conditions of the experiments comply with those of PRI-ANAC [10]. The tournaments are run on GENIUS platform (version

Table 1: Average individual utilities of agents

Selection algorithm	<i>UCB(0.01)</i>	<i>UCB(0.05)</i>	<i>UCB(0.1)</i>	<i>UCB(0.5)</i>	<i>UCB(1)</i>	<i>EG(0)</i>	<i>EG(0.1)</i>	<i>EG(0.2)</i>	<i>EG(1)</i>	<i>pure-MAB</i>
Meta-Agent	0.7788	0.7765	0.7725	0.7382	0.7201	0.7787	0.7706	0.7631	0.6979	0.7406
<i>Atlas3</i>	0.7444	0.7443	0.7451	0.7474	0.7477	0.7441	0.7445	0.7451	0.7475	0.7501
<i>CaduceusDC16</i>	0.7138	0.7135	0.7134	0.7119	0.7104	0.7137	0.7130	0.7127	0.7090	0.7133
<i>kawaii</i>	0.7305	0.7304	0.7304	0.7274	0.7254	0.7312	0.7299	0.7293	0.7221	0.7265
<i>ParsCat</i>	0.6867	0.6872	0.6877	0.6875	0.6869	0.6869	0.6862	0.6862	0.6863	0.6881
<i>Rubick</i>	0.6658	0.6664	0.6652	0.6652	0.6648	0.6658	0.6654	0.6646	0.6644	0.6717
<i>YXAgent</i>	0.7132	0.7129	0.7121	0.7050	0.7006	0.7130	0.7110	0.7097	0.6944	0.6994

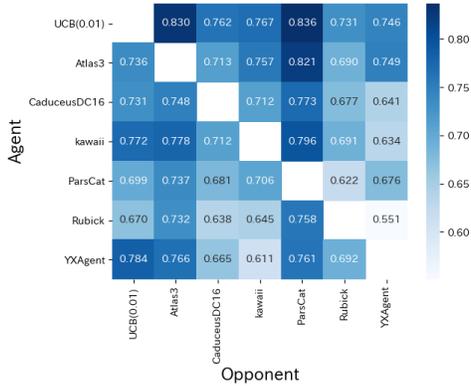


Figure 1: Average individual utilities of agents for each opponent

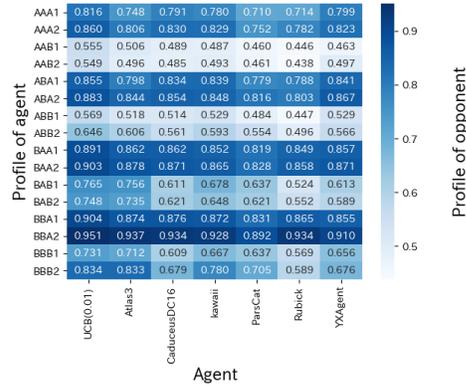


Figure 2: Average individual utilities of agents for each its own profile

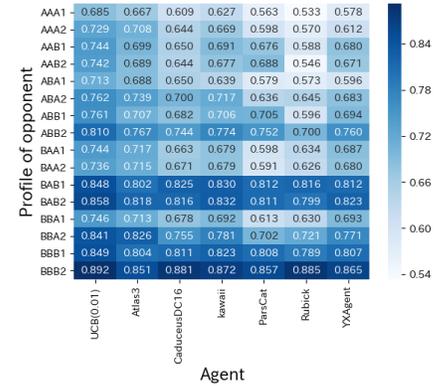


Figure 3: Average individual utilities of agents for each opponent's profile

9.1.1) [13]. The tournaments comprise seven agents (Our Agent (or pure-MAB), Atlas3, CaduceusDC16, kawaii, ParsCat, Rubick and YXAgent). We use the domain used in PRIANAC [10]. The domain comprises 16 profiles and includes various types of scenarios such as cooperative, competitive, and unfair. Each negotiation session comprises of two different agents. Sessions in all combinations of agents and profiles are conducted. Each session repeats 100 times with the same conditions (5,400 combinations and 540,000 negotiations). The negotiation protocol is alternating offers protocol [15] and the deadline of each negotiation is 10 seconds. These settings are common knowledge of the agents. We compare 10 agents with different algorithms and parameters: *UCB(0.01)*: UCB algorithm with $c = 0.01$; *UCB(0.05)*: UCB algorithm with $c = 0.05$; *UCB(0.1)*: UCB algorithm with $c = 0.1$; *UCB(0.5)*: UCB algorithm with $c = 0.5$; *UCB(1)*: UCB algorithm with $c = 1$; *EG(0)*: ϵ -greedy with $\epsilon = 0$ (greedy algorithm); *EG(0.1)*: ϵ -greedy with $\epsilon = 0.1$; *EG(0.2)*: ϵ -greedy with $\epsilon = 0.2$; *EG(1)*: ϵ -greedy with $\epsilon = 1$ (random selection); *pure-MAB*: pure-MAB agent by Ilany and Gal [9].

Table 1 shows the average individual utility of each agent. A column describes the results of each tournament which comprises of our agent and existing agents. *UCB(0.01)*, *UCB(0.05)*, *UCB(0.1)*, *EG(0)*, *EG(0.1)*, and *EG(0.2)* significantly outperform all other agents (Mann-Whitney U-test, $p < 0.01$). Our meta-strategy with an appropriate parameter works effectively in multi-time negotiation. The utility of “Oracle agent”, which selects optimal strategy by computing in retrospect, is 0.7847. Our agents reached 99% performance of “Oracle agent” in individual utility. As c and ϵ get smaller,

the individual utility of our agents become higher. It implies that exploring is not important for the strategy selection of our meta-strategy. Agents rarely overestimate or underestimate the performance of negotiation strategy because the variance of individual utilities is small in the negotiations in the same situation. Thus, agents can find the best strategy for the situation with a small amount of exploration.

Figure 1 - 3 show the average individual utility of an agent for each opponent, each its own profile, and opponent’s profile in the tournaments for agent *UCB(0.01)*. They demonstrate that individual utility of the agent depends on the opponent’s strategy, its own profile and opponent’s profile. Our agent has a higher individual utility than all other agents in various condition although our agent adopts these agents’ negotiation strategies.

4 CONCLUSION AND FUTURE WORK

In this paper, we proposed a meta-strategy that provides a suitable negotiation strategy for the situation to the agent for multi-time negotiation. One of the possible future research is to consider dynamic conditions, such as changing the preferences through repetitions, or if the opponent adopts a meta-strategy. Although the exploration was not important under the conditions in this paper, the effective exploration of the individual negotiation strategy is necessary in more dynamic conditions.

REFERENCES

- [1] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. 2002. Finite-time Analysis of the Multiarmed Bandit Problem. *Machine Learning* 47, 2-3 (2002), 235–256. <https://doi.org/10.1023/A:1013689704352>
- [2] Reyhan Aydogan, Tim Baarslag, Katsuhide Fujita, Koen Hindriks, Takayuki Ito, and Catholijn Jonker. 2016. The Seventh International Automated Negotiating Agents Competition (ANAC2016). <http://web.tuat.ac.jp/katfuji/ANAC2016/>. (2016).
- [3] Reyhan Aydogan, Tim Baarslag, Katsuhide Fujita, Takayuki Ito, Dave de Jonge, Catholijn Jonker, and Johnathan Mell. 2017. The Eighth International Automated Negotiating Agents Competition (ANAC2017). <http://web.tuat.ac.jp/katfuji/ANAC2017/>. (2017).
- [4] Reyhan Aydogan, Tim Baarslag, Katsuhide Fujita, Takayuki Ito, Catholijn Jonker, Dave de Jonge, and Johnathan Mell. 2018. The Ninth International Automated Negotiating Agents Competition (ANAC2018). <http://web.tuat.ac.jp/katfuji/ANAC2018/>. (2018).
- [5] Tim Baarslag, Reyhan Aydogan, Koen V. Hindriks, Katsuhide Fujita, Takayuki Ito, and Catholijn M. Jonker. 2015. The Automated Negotiating Agents Competition, 2010-2015. *AI Magazine* 36, 4 (2015), 115–118. <http://www.aaai.org/ojs/index.php/aimagazine/article/view/2609>
- [6] Tim Baarslag, Katsuhide Fujita, Enrico Gerding, Koen Hindriks, Takayuki Ito, Nick R. Jennings, Catholijn Jonker, Sarit Kraus, Raz Lin, Valentin Robu, and Colin Williams. 2013. Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. *Artificial Intelligence Journal (AIJ)* 198 (2013), 73–103.
- [7] Tim Baarslag, Mark Hendriks, Koen V. Hindriks, and Catholijn M. Jonker. 2013. Predicting the Performance of Opponent Models in Automated Negotiation. In *2013 IEEE/WIC/ACM International Conferences on Intelligent Agent Technology, IAT 2013, 17-20 November 2013, Atlanta, Georgia, USA*. IEEE Computer Society, 59–66. <https://doi.org/10.1109/WI-IAT.2013.91>
- [8] Shaheen Fatima, Sarit Kraus, and Michael Wooldridge. 2014. *Principles of Automated Negotiation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511751691>
- [9] Litan Ilany and Ya'akov Gal. 2016. Algorithm selection in bilateral negotiation. *Autonomous Agents and Multi-Agent Systems* 30, 4 (01 Jul 2016), 697–723. <https://doi.org/10.1007/s10458-015-9302-8>
- [10] Takayuki Ito, Catholijn Jonker, Reyhan Aydogan, Tim Baarslag, Katsuhide Fujita, Satoshi Morinaga, Takashi Yoshida, and Yasser Mohammad. 2018. Pacific Rim International Automated Negotiation Agents Competition (PRIANAC). <http://web.tuat.ac.jp/katfuji/PRIANAC2018/>. (2018).
- [11] Nicholas R Jennings, Peyman Faratin, Alessio R Lomuscio, Simon Parsons, Michael J Wooldridge, and Carles Sierra. 2001. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation* 10, 2 (2001), 199–215.
- [12] Sarit Kraus. 2001. *Strategic Negotiation in Multiagent Environments*. MIT Press. <http://mitpress.mit.edu/books/strategic-negotiation-multiagent-environments>
- [13] Raz Lin, Sarit Kraus, Tim Baarslag, Dmytro Tykhonov, Koen Hindriks, and Catholijn M. Jonker. 2014. Genius: An Integrated Environment for Supporting the Design of Generic Automated Negotiators. *Computational Intelligence* 30, 1 (2014), 48–70. <https://doi.org/10.1111/j.1467-8640.2012.00463.x>
- [14] Herbert Robbins. 1952. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.* 58, 5 (1952), 527–535.
- [15] Ariel Rubinstein. 1982. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society* (1982), 97–109.