A Scalable Opponent Model Using Bayesian Learning for Automated Bilateral Multi-issue Negotiation

Extended Abstract

Shengbo Chang
Tokyo University of Agriculture and Technology
Tokyo, Japan
s215450r@st.go.tuat.ac.jp

Katsuhide Fujita
Tokyo University of Agriculture and Technology
Tokyo, Japan
katfuji@cc.tuat.ac.jp

ABSTRACT

Learning an opponent's preference is critical to achieving a win-win situation in automated bilateral multi-issue negotiations. Most of the existing opponent preference-learning techniques are not scalable to many kinds of opponents with different strategies due to their strong assumptions on an opponent's concession pattern. This study enables a more general assumption into the Bayesian-learning-based opponent model to address the mentioned disadvantage. The proposed method is experimentally compared with state-of-the-art opponent models and found to have higher accuracy and greater scalability in most cases.

KEYWORDS

Multi-Issue negotiation; Opponent Modeling; Bayesian Learning

ACM Reference Format:

Shengbo Chang and Katsuhide Fujita. 2023. A Scalable Opponent Model Using Bayesian Learning for Automated Bilateral Multi-issue Negotiation: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

In automated multi-issue negotiations, agents uses so called opponent models to learn and estimate an opponent's preference profile (usually a linear additive utility function maps each offer with a real value in $[0, 1]$, see Eq. 1) from exchanged offers [2]. These models are distinct from traditional learning methods in that the labels of training examples (i.e., the utility values of an opponent's offers) are usually unavailable. This has led most existing opponent models to be based on one or more assumptions about an opponent's behavior, such as the assumption that the issues proposed more frequently have higher preferences, that the issues with values that change less are more important, and that the opponent follows a specific concession function [1]. However, these strong assumptions limit the scalability of the models to the opponents of different behavior pattern.

$$U(\omega) = \sum_{i=1}^{N} w_{I_i} e_{I_i}(\omega[I_i]) \tag{1}$$

where w_{I_i} is the weight of an issue I_i ($\sum_{i=1}^{N} w_i = 1$), and $e_{I_i}(\omega[I_i])$ is an evaluation function mapping the value of I_i in the outcome ω to a real number normalized to the $[0, 1]$ range.

2 SCALABLE OPPONENT MODEL

This study improves the Bayesian-learning-based opponent model by incorporating a more general assumption that "the opponent tends to make concession over time". The model estimates the opponent's utility function $\hat{U}_t$ by the expectation of a set of utility function hypotheses (Eq.2), each with a probability of being the opponent's true utility function under the bidding sequence B_t up to time t (Bayes' rule, Eq.3). The likelihood $P(B_t|h_k)$ is recursively updated with the conditional probability $P(b_t|h_k, B_{t-1})$ (see Eq. 4). Eq. 5 of calculating $P(b_t|h_k, B_{t-1})$ realizes the assumption by assigning a lower probability to hypotheses suggesting that the current bid b_t has a higher utility than the previous bid b_{t-1}, with the probability decreasing as the amount of support increases. Conversely, it assigns a higher probability to hypotheses supporting that b_t has the same or lower utility than b_{t-1} (see Figure. 1).

$$\hat{U}_t = \sum_{i=1}^{|H|} P_t(h_i) h_i \tag{2}$$

where H is the hypothesis space; $\hat{U}$ is the estimated utility function.

$$P_t(h_k) = P(h_k|B_t) = \frac{P(h_k)P(B_t|h_k)}{P(B_t)} \tag{3}$$

$$P(B_t|h_k) = P(b_t, B_{t-1}|h_k) = P(b_t|h_k, B_{t-1})P(B_{t-1}|h_k) \tag{4}$$

where $P(h_k)$ is the prior probability of hypothesis h_k.

$$P(b_t|h_k, B_{t-1}) = P(b_t|h_k, b_{t-1}) = \frac{1}{\lambda\sigma\sqrt{2\pi}} e^{-\frac{\delta}{2\sigma^2}}$$
$$\delta = \begin{cases} h_k(b_t) - h_k(b_{t-1}), & if \;\; h_k(b_t) - h_k(b_{t-1}) > 0 \\ 0, & otherwise \end{cases} \tag{5}$$

To ensure scalability across domains of varying sizes, we adopt the hypothesis space and evaluation method from the scalable Bayesian learning opponent model [4]. This model discretizes the continuous utility function space into rankings and distinguishes between weight hypotheses and value hypotheses. Additionally, to reduce computational costs, it calculates the likelihood probability of a hypothesis based on its probability conditioned on the mean of all other hypotheses.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

Table 1: Accuracy versus different opponent agents. The proposed scalable model, CUHKagent value model, Hardheaded frequency model, and existing Bayesian learning model as *Proposed*, *CUHKagent*, *Hardheaded*, and *Bayesian*, respectively.

Agent category	Strategy parameter	Pearson correlation (whole)				Pearson correlation (the last 250 rounds)			
		Proposed	Bayesian	CUHKagent	Hardheaded	Proposed	Bayesian	CUHKagent	Hardheaded
Time	$\alpha = 0.1$	**0.78**	-0.30	-0.35	-0.48	**0.77**	-0.32	-0.38	-0.33
	$\alpha = 1.0$	**0.77**	0.47	**0.77**	0.59	**0.84**	0.45	0.78	0.30
	$\alpha = 10.0$	0.66	0.47	**0.69**	0.36	**0.78**	0.49	0.77	0.55
Offset	$P_{max} = 0.7$	**0.63**	0.28	0.60	0.12	**0.64**	0.26	0.57	-0.18
	$P_{max} = 0.8$	**0.75**	0.35	0.73	0.34	**0.77**	0.33	0.72	-0.13
	$P_{max} = 0.9$	0.75	0.44	**0.78**	0.52	0.77	0.42	**0.78**	-0.00
Reservation	$R = 0.3$	0.73	0.47	**0.77**	0.61	**0.83**	0.47	0.78	0.70
	$R = 0.5$	0.67	0.48	**0.76**	0.56	0.74	0.48	**0.78**	0.74
	$R = 0.7$	0.64	0.48	**0.75**	0.47	0.66	0.48	**0.78**	0.62

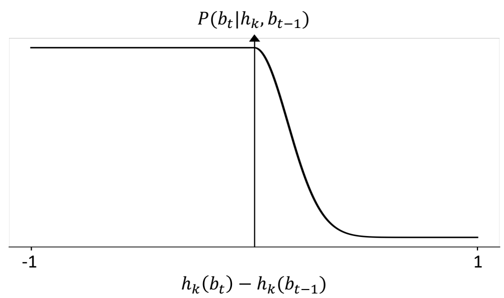

Figure 1: Conditional probability distribution of the behavior assumption

Table 2: Agents used for the bidding trace generation

Agent type	Parameters	Noise
Time-dependent agents	$\alpha \in \{0.1, 1.0, 10.0\}$	$\sigma \in \{0,$
Offset agents	$P_{max} \in \{0.7, 0.8, 0.9\}$	0.005,
Reserved agents	$R \in \{0.3, 0.5, 0.7\}$	0.05$\}$

3 EXPERIMENTS

We compared the proposed model with three state-of-the-art opponent models: CUHKagent value model [3], Hardheaded frequency model [7], and scalable Bayesian model [4]. Based on relevant papers and GENIUS codes [5], we implemented the opponent models ourselves using Python and NegMAS [6] and ensured that each opponent model processed the same amount of data. We used the *Pearson correlation of bids* to evaluate the opponent models, which has been proved to be efficient [1].

We used the same negotiation scenarios with [1]. The length of the bidding trace is set to $t_{max} = 5000$. We recorded the bidding traces of 27 agents that were the Cartesian products of nine different strategies and three different noise levels. Table 2 lists the 27 agents: the time-dependent agents selected bid according to a target utility of $u_t = 1 - (t/t_{max})^{\alpha}$; The offset agents were time-dependent and did not start from their best bid, their target utility followed $u_t = P_{max} \cdot (1 - t^{\alpha})$; the reservation agents were time-dependent agents with a reservation value, and their target utility followed $u_t = 1 - (1 - R) \cdot t^{\alpha}$; the noise is added to the target utility following $u_t = u_t + u_{noise}$, where $u_{noise} \sim N(0, \sigma^2)$, which is used to account for agents that may not follow only the time-dependent strategies.

Table 1 presents the average accuracy results for each opponent agent across the entire negotiation and at the end of the negotiation. The proposed model only outperformed the CUHKagent value model in a few cases in terms of accuracy over the entire negotiation. However, the CUHKagent value model performed poorly with an accuracy value of -0.35 for the time-dependent agent with $\alpha = 0.1$. In contrast, the proposed model reached an accuracy of 0.64 even in the worst case against reservation agents with $R = 0.7$. Against offset agents, our model showed greater scalability to different starting utilities of the opponent. These results demonstrate that the proposed model can perform more robustly than existing models, which may perform poorly in some cases.

In addition, we believe that the opponent model's accuracy at the end of a negotiation is more critical for an agent than that at the start of the negotiation, especially when encountering stubborn opponents in real negotiations. This is because it is often wise to wait until the opponent makes more concessions. In terms of accuracy at the end of the negotiation, the proposed opponent model generally showed higher accuracy than the CUHKagent value model. Comparing accuracy over the entire negotiation to that at the end, the proposed model's accuracy increased more than that of the state-of-the-art models, especially for the time-dependent agent with $\alpha = 10.0$. One possible reason for this finding is that if the opponent makes more concessions in a step, there may be greater differences in the values between the two bids, providing our model with more information to improve accuracy.

4 CONCLUSION AND FUTURE WORK

Learning an opponent's preference was critical to achieving a win-win situation in automated bilateral multi-issue negotiations. The main idea of this study was to realize an opponent model that is scalable to different opponents by implementing a more general behavioral assumption into the Bayesian learning model. We demonstrated that the proposed method had higher accuracy and greater scalability in most cases.

One possible future direction is applying the proposed model to multilateral multi-issue negotiations.

ACKNOWLEDGMENTS

Thanks Dr. Baarslag for providing the necessary scenario files of experiments and more detailed information on the accuracy measure. This work was supported by JSPS KAKENHI Grant Numbers 22H03641, 19H04216, and JST FOREST (Fusion Oriented REsearch for disruptive Science and Technology) Grant Number JPMJFR216S.

REFERENCES

[1] Tim Baarslag, Mark Hendrikx, Koen Hindriks, and Catholijn Jonker. 2013. Predicting the Performance of Opponent Models in Automated Negotiation. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, Vol. 2. IEEE, 59–66. https://doi.org/10.1109/WI-IAT.2013.91

[2] Shaheen Fatima, Sarit Kraus, and Michael Wooldridge. 2014. *Principles of Automated Negotiation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511751691

[3] Jianye Hao and Ho-fung Leung. 2014. CUHKAgent: An Adaptive Negotiation Strategy for Bilateral Negotiations over Multiple Items. In *Novel Insights in Agent-based Complex Automated Negotiation*, Ivan Marsa-Maestre, Miguel A. Lopez-Carmona, Takayuki Ito, Minjie Zhang, Quan Bai, and Katsuhide Fujita (Eds.). Springer Japan, Tokyo, 171–179. https://doi.org/10.1007/978-4-431-54758-7_11

[4] Koen Hindriks and Dmytro Tykhonov. 2008. Opponent modelling in automated multi-issue negotiation using Bayesian learing. In *AAMAS 2008*, L Padgham, D Parkes, J Muller, and S Parsons (Eds.). Lightning Source UK Ltd., 331–338. null ; Conference date: 12-05-2008 Through 16-05-2008.

[5] 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

[6] Yasser Mohammad, Shinji Nakadai, and Amy Greenwald. 2020. NegMAS: A Platform for Automated Negotiations. In *PRIMA 2020: Principles and Practice of Multi-Agent Systems: 23rd International Conference, Nagoya, Japan, November 18–20, 2020, Proceedings* (Nagoya, Japan). Springer-Verlag, Berlin, Heidelberg, 343–351. https://doi.org/10.1007/978-3-030-69322-0_23

[7] Thijs van Krimpen, Daphne Looije, and Siamak Hajizadeh. 2013. HardHeaded. In *Complex Automated Negotiations: Theories, Models, and Software Competitions*, Takayuki Ito, Minjie Zhang, Valentin Robu, and Tokuro Matsuo (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 223–227. https://doi.org/10.1007/978-3-642-30737-9_17