Opponent’s Preference Estimation Considering Their Offer Transition in Multi-Issue Closed Negotiations

Yuta HOSOKAWA (a), Nonmember and Katsuhide FUJITA (b), Member

SUMMARY In recent years, agreement technologies have garnered interest among agents in the field of multi-agent systems. Automated negotiation is one of the agreement technologies, in which agents negotiate with each other to make an agreement so that they can solve conflicts between their preferences. Although most agents keep their own preferences private, it is necessary to estimate the opponent’s preferences to obtain a better agreement. Therefore, opponent modeling is one of the most important elements in automated negotiating strategy. A frequency model is widely used for opponent modeling because of its robustness against various types of strategy while being easy to implement. However, existing frequency models do not consider the opponent’s proposal speed and the transition of offers. This study proposes a novel frequency model that considers the opponent’s behavior using two main elements: the offer ratio and the weighting function. The offer ratio stabilizes the model against changes in the opponent’s offering speed, whereas the weighting function takes the opponent’s concession into account. The two experiments conducted herein show that our proposed model is more accurate than other frequency models. Additionally, we find that the agent with the proposed model performs with a significantly higher utility value in negotiations.

key words: multi-agent system, bilateral multi-issue negotiation, negotiation strategy, opponent preferences estimation, opponent modeling

1. Introduction

In recent years, a growing interest in agreement technologies has been observed among agents in the field of multi-agent systems. Autonomous agents follow their owner’s objective or preference and act accordingly. Thus, agents often experience conflicts with each other. Automated negotiation is attracting attention as one method of solving these problems. Autonomous agents negotiate with each other to achieve a better agreement that resolves their conflicts in automated negotiation [1], [2].

The Automated Negotiating Agents Competition (ANAC) has been held since 2010 to develop research on automated negotiation ([3], [4] etc.). Many realistic settings, such as the alternating-offers protocol and time discounting, are applied to ANAC to explore the trends of agent strategies and negotiating protocols. The Pacific Rim International Automated Negotiation Agents Competition (PRI-ANAC), which is the latest negotiating competition, focuses on multi-time negotiation [5]. In the setting of multi-time negotiation, agents repeatedly negotiate under the same environment, which commonly occurs in business scenarios. In such a setting, the history of past negotiations must be efficiently utilized.

Most agents keep their own preference information secret to avoid exploitation by the opponent. However, for the agents, deciding on their behavior and finding an optimal agreement without any knowledge about the opponent’s preferences is difficult. Accordingly, constructing the opponent’s preference model during a negotiation is an effective approach for making a better agreement [6]. The frequency model is one of the most popular methods for opponent modeling; it assumes that the opponent steadily concedes from higher to lower utility and builds an estimation based on the frequency of the opponent’s proposals. The frequency model can be widely applied against various opponent strategies because of its weak assumption. Many existing frequency models adopt raw numbers of frequency for preference estimation. Accordingly, the results might be changed by the opponent’s offering speed or the negotiation deadline. Furthermore, many frequency models do not consider the time series information of the opponent’s offers because they treat all offers uniformly.

Herein, we propose a novel frequency modeling method, called the time-dependent ratio model. The main elements of the model are offer ratio and weighting function, which pinpoints the opponent’s behavior. In the proposed model, the offer ratio stabilizes the model against changes in the opponent’s offering speed, and the weighting function considers the opponent’s concession in the negotiation. We performed two experiments to compare the models: the first experiment analyzes the estimation accuracy, whereas the second one measures model performance. The experimental results show the accuracy of our proposed model. The results also demonstrate that our model outperforms other models in negotiations. We designed our model to be applicable to a single-shot negotiation, in which agents have no prior information about the opponent or the previous negotiations, and evaluated in such settings. However, we think that our model might easily adopt to multi-time negotiations.

The remainder of the paper is organized as follows: Sect. 2 describes the negotiation settings in this work; Sect. 3
introduces two existing frequency-based opponent models; Sect. 4 presents the proposed opponent modeling method; Sect. 5 discusses the experimental analyses; Sect. 6 introduces related work on opponent modeling; and finally, Sect. 7 concludes the paper.

2. Bilateral Multi-Issue Negotiation

In this work, we focus on bilateral multi-issue negotiations, in which two agents aim to make an agreement maximizing their own utility. Agents exchange offers (bids) with each other using the alternating-offers protocol. Each agent decides if the previous offer from the opponent is acceptable in its turn. The agent makes a new offer when it does not accept the opponent’s offer. Each negotiation has a common deadline, specified in real-time or by a number of rounds. Agents can obtain the normalized time $t$ in the negotiation; the negotiation starts at $t = 0$ as well as it ends at $t = 1$.

The negotiation domain comprises $n$ issues $[I_1, I_2, \ldots, I_n]$. Each issue $I_i$ contains $K_i$ values (options) $[v_{i1}, v_{i2}, \ldots, v_{ik}]$. An offer is represented by a vector $\vec{b} = [v_{i1}, v_{i2}, \ldots, v_{in}]$, containing one option for each issue.

Every agent has a unique utility function that represents its preferences. The utility function $U(\vec{b})$ maps a bid $\vec{b}$ to a utility which is a real number in the range between 0 and 1. We assume that the utility function is linearly additive. In this case, the preferences are divided into issue weight $w_i$ for each issue and evaluation value $eval(v_{ic})$ for each option. The utility for a bid $\vec{b} = [v_{i1}, v_{i2}, \ldots, v_{in}]$ is defined as follows:

$$U(\vec{b}) = \sum_{i=1}^{n} w_i \times eval(v_{ic}).$$

It is assumed that $\sum_{i=1}^{n} w_i = 1$ and $w_i \geq 0$ for any $i$, and $eval(v_{ic})$ is assumed to be a real number in the range between 0 and 1 for any $i$ and $c$. Note that both agents receive utilities of 0 when the deadline arrives without any agreement.

The preference of each agent is not open to the other. However, an agent can make a better agreement when it estimates the opponent’s preference during a negotiation [7].

3. Frequency Models for Estimating the Opponent’s Preferences

Frequency models are widely used for estimating the opponent’s preferences in automated negotiation. Although frequency models are simple, they are known to be more robust against various opponent strategies [7], [8].

3.1 HardHeaded Agent’s Frequency Model

Multiple types of frequency-based modeling methods have been proposed. The HardHeaded agent’s frequency model [9] is known as the most basic and popular method among these methods. This model makes the following assumptions about the opponent’s strategy.

- The opponent steadily reduces the utility of the offers proposed.
- The opponent prefers to explore the negotiation domain widely rather than make the same offer repeatedly.
- The opponent tends to change the choices on less-favored issues than on the most preferred issue.

This model estimates the issue weights and evaluation values separately. Estimation of the evaluation values is based on the frequency of appearance of each option in the opponent’s offers. The estimated values are defined as the normalized frequency of selection.

The issue weights are estimated based on the frequency of the changes between two consecutive offers. Each issue weight is considered to be the same at the beginning of the negotiation. If five issues exist in the domain, initially, each issue has a weight of 0.2. When a new offer is made by the opponent, the model verifies whether the value of an issue is changed. If the issue selection does not change, the model increases the weight of the issue because that issue is considered to be more preferred by the opponent.

This frequency model can make an effective estimation against some specific agents. However, many agents’ strategies do not completely follow the assumptions, and the gap between the assumptions and the opponent’s strategy leads to low accuracy. First, agents tend to repeat the same offer over and over again to avoid revealing much information. In that case, the estimated value of the specific option increases because the option is frequently offered. Second, some agents prefer to stochastically explore neighboring bids rather than strictly follow their concession rules. The HardHeaded agent’s frequency model only considers a pair of consecutive offers; hence, it misses the global trend of concession of such agents.

3.2 Distribution-Based Frequency Model

The distribution-based frequency model [10] is the extended opponent modeling method of previous frequency models. Similar to the HardHeaded agent’s frequency model, this model also estimates the issue weights and evaluation values independently.

The estimation of the evaluation values is based on the following assumption: the more preferred is the value, the more frequently bids are offered, which is identical to the assumption in the HardHeaded agent’s frequency model. However, instead of simply normalizing the frequency of offers, the distribution-based frequency model calculates the estimated value $eval(v_{ic})$ as Eq. (2)

$$\bar{eval}(v_{ic}) = \frac{1 + \text{freq}(v_{ic})}{\max_{k \in [1,2,\ldots,K_i]} (1 + \text{freq}(v_{ik}))},$$

where $\text{freq}(v_{ic})$ is the frequency count of the value $v_{ic}$ in the issue $I_i$, and $K_i$ is the number of options for the issue $I_i$.

Laplace smoothing is applied for this estimation function, resulting in smoothed distributions at the beginning of
the negotiation. This method also provides some evaluation values to options that are never offered.

The smoothed frequencies are passed by the exponential function of $\gamma$, where $\gamma$ is a constant value in $0 \leq \gamma \leq 1$. When the opponent makes the same set of offers over and over again, only the frequency of the options in the offers will increase to a greater extent. The exponential filter slows the growth of such imbalanced values. They found out that it is effective when $\gamma$ was set to 0.25; thus we use that value in our experiments.

As mentioned in the previous section, estimating the issue weights of agents that stochastically explore the negotiation space is difficult when only pairs of consecutive offers are used. The distribution-based frequency model estimates the issue weights by dividing the consecutive offers into a fixed length of windows for solving the above problem.

Each window contains a fixed number of consecutive offer bids from the opponent. The model tests whether the distribution of selection for each issue has been changed for estimating the issue weights. Comparing the window distributions makes the estimation more robust against stochastically exploring agents. Furthermore, the model can observe the opponent’s general trend of concession through these comparisons.

4. Time-Dependent Ratio Model

4.1 Evaluation Value Estimation

Each time a negotiating agent takes its turn, it runs a calculation, such as opponent modeling, making a counteroffer, or deciding whether the offer is acceptable. Consequently, agents take different amounts of time to offer a new bid depending on the situation. On the other hand, most of the agents employ time-based concession strategies, in which the agents determine the proposing bids according to the normalized time. (See Sect. 4.3.) Therefore, agents may send different numbers of similar offers within a given time period. However, many existing opponent models directly use the value of frequency counts. For example, the estimation results of the distribution-based frequency model in Sect. 3.2 are different, as shown in Table 1. Additionally, the opponent can lower the estimation accuracy by changing the offering speed even if the agent follows a concession function. Furthermore, if the negotiation deadline is defined by a round, the estimation accuracy will depend on the fixed deadline.

Many existing agents concede as time goes on. Such conceding agents select the most preferred value in the beginning, and other values are not offered. As the negotiation time passes, less preferred values are offered as compromising plans. Many frequency models treat all proposals uniformly without considering such offer changes.

For these reasons, our proposed model, called the time-dependent ratio (TDR) model, uses an offer ratio that is not varied by the offering speed or negotiation deadline and applies a weight function.

| Table 1 | Estimation results against different-speed opponents. $\gamma$ is set to 0.25. |
|---------|-----------------------------------------------------------------------|
|         | value $v_1$ | value $v_2$ | value $v_3$ |
| freq($v_j$) | 40          | 0            | 20          |
| $\left(1 + \text{freq}(v_j)\right)^{\gamma}$ | 2.53        | 1.00         | 2.14        |
| eval($v_j$) | 1.00        | 0.40         | 0.85        |

(a) Estimation against an agent having slow offering speed

|         | value $v_1$ | value $v_2$ | value $v_3$ |
|---------|-------------|-------------|-------------|
| freq($v_j$) | 4000        | 0           | 2000        |
| $\left(1 + \text{freq}(v_j)\right)^{\gamma}$ | 7.95        | 1.00        | 6.69        |
| eval($v_j$) | 1.00        | 0.13        | 0.84        |

(b) Estimation against an agent having fast offering speed

Offer ratio $P(v_j, t)$

Instead of the frequency counts, which many frequency models use, our model uses the offer ratio for each time period. The number of divisions of time must be large enough to capture the concessions of the opponent agent, but small enough that the count within the time period is not sparse. For that reason, the TDR model separates the negotiation time into 100 periods and records the bids given by the opponent for each time period. However, since the number of 100 has set by a heuristics, further experiments to improve its accuracy might be needed. $T = \{0.01, 0.02, \ldots, 1.0\}$ denotes the set of separated time periods, where $t \in T$ means the time period of $(t - 0.01, t]$. We define the offer ratio $P(v_j, t)$ as the ratio of selections of the value $v_j$ in issue $I_t$ at time $t$. The TDR model invariably maintains the estimation results as is, as long as the concession function of the opponent remains the same.

Weight depends on the time of offering $w(t)$

Offers from the opponent that follow the concession function change as time passes. We weight the offer ratio with a time-dependent function $w(t)$ to consider such changes.

We define the estimation of the evaluation value $\text{eval}(v_j)$ as Eq. (3) using the offer ratio $P(v_j, t)$ and the weighting function $w(t)$.

$$\text{eval}(v_j) = \max_{k \in \{1, 2, \ldots, K_t\}} \left[ \sum_{t \in T} w(t) \times P(v_j, t) \right]^{\gamma}$$

Note that $K_t$ is the number of options for the issue $I_t$, and $t$ in summation function moves through the separated time periods from 0.01 to 1.0. We use the weighted sum of the offer ratio instead of the frequency in the distribution-based frequency model in Sect. 3.2. Laplace smoothing can provide some evaluation values for less-selected options. Furthermore, the exponential filter of $\gamma$ can fix the unbalanced values of the options offered over and over again.
4.2 Issue Weight Estimation

We use the algorithm of the distribution-based frequency model in Sect. 3.2 to estimate the issue weights, but we change the window size of a fixed number of opponent’s offers to the opponent’s offers in the time periods.

As mentioned in Sect. 4.1, the computational time to offer a new bid is different for each agent. In addition, the opponent has the liberty to intentionally change the offering speed. The deadline round affects the total number of exchanged offers if the deadline is specified by the number of rounds. Accordingly, similar to the estimation of the evaluation values, we separate the bid history according to the time period. The total number of windows during a negotiation is set as 100.

4.3 Determining the Parameters

We analyzed the negotiation strategies of ANAC agents and ran a preliminary experiment to settle \( w(t) \) and \( \gamma \) in Eq. (3). We selected the top five agents of the ANAC2015 [11] and ANAC2016 [12] to analyze the concession strategies. We found that all of the ten agents used time-based concession strategies and most of them employed the concession functions in the form of Eq. (4).

\[
U_{\text{target}}(t) = 1 - (1 - U_{\text{min}}) \times t^{1/\gamma} \tag{4}
\]

Many agents change the concession parameters depending on the negotiation domains and the behavior of the opponents. The ten agents we selected and their detailed concession strategies are as follows:

- **Atlas3**: Follows Eq. (4). \( e \) is always set to 1 and \( U_{\text{min}} \) is varied by the estimated expected utility value and the domain settings.
- **ParsAgent**: Follows Eq. (4). \( e \) is changed between 0.15 and 0.2, and \( U_{\text{min}} = 0 \). However, when \( U_{\text{target}}(t) \) is less than 0.7, the agent considers it as a constant value of 0.7.
- **RandomDance**: Follows Eq. (4). \( e \) is changed in \( e \geq 1/3 \), and \( U_{\text{min}} \) is varied by the negotiation domain and the behavior of the opponent.
- **Kawaii**: Follows Eq. (4). \( e = 0.5 \) and \( U_{\text{min}} = 0.8 \).
- **AgentBuyog**: Follows Eq. (4). \( e = 0.56 \) and \( U_{\text{min}} \) is varied depending on the behavior of its opponent.
- **Caduceus**: This agent selects one of the concession functions of Atlas3, ParsAgent, RandomDance, Kawaii, and its original linear function (\( e = 1.0 \)) depending on the situation.
- **YXAgent**: \( U_{\text{target}}(t) \) is always set to 0.9 regardless of the opponent’s behavior or the negotiation domain.
- **ParsCat**: This agent uses a complex time-based concession function to make it difficult for the opponent to estimate its utility function. Although its concession function is not decreasing monotonically, the global trend of the function is to decrease over time.

- **Farma**: Follows Eq. (4). It changes \( e \) and \( U_{\text{min}} \) depending on the opponent’s behavior.
- **MyAgent**: Follows Eq. (4). \( e \) is varied between 0.01 and 1.65 based on the opponent’s offers, and \( U_{\text{min}} = 0 \).

As clarified above, agents may change their behavior, such as concession strategy, depending on the domain settings and their opponents’ behavior. Therefore, we performed negotiations with all combinations of the ten agents in various domains and recorded the relationship between the agents’ utility functions and their offered bid sequences. The negotiation deadline was set to 60 seconds in real-time, and we measured the number of bids offered by each agent per second. In total, We ran 208,350 negotiations with using 30 kinds of domains available in Genius platform [13].

In these negotiations, the number of bids sent per second varied from less than 1 to more than 3000, and the average number was 392. These significant differences in speed were similar for all agents, so we can observe that the offering speed varies greatly depending on the situation. To determine the parameters \( w(t) \) and \( \gamma \) in our model, we conducted a grid search with inputting the bid sequences to our model. The candidates for \( w(t) \) were a constant (\( w(t) = 1 \)), decreasing functions (\( w(t) = 1 - t, 1 - t, 1 - t^{1/3} \)), and increasing functions (\( w(t) = t, t, t^{1/3} \)), and the candidates for \( \gamma \) were \( 1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \). Table 2 shows the average difference between the real evaluation values and the estimated evaluation values. From these results, \( w(t) = 1 - t^{1/3} \) and \( \gamma = \frac{1}{3} \) were determined.

### Table 2 Average difference between evaluation values

| \( w(t) \) \& \( \gamma \) | \( \frac{1}{4} \) | \( \frac{1}{3} \) | \( \frac{1}{2} \) | \( \frac{3}{4} \) | \( \frac{1}{3} \) | \( \frac{1}{5} \) |
|-----------------|------|------|------|------|------|------|
| 1               | 0.2940 | 0.2082 | 0.1778 | 0.1743 | 0.1825 | 0.1935 |
| 1 - \( t^{1/3} \) | 0.2971 | 0.2081 | 0.1763 | 0.1730 | 0.1814 | 0.1922 |
| 1 - \( t \)     | 0.3007 | 0.2032 | 0.1733 | 0.1728 | 0.1821 | 0.1948 |
| 1 - \( t^{1/3} \) | 0.3033 | 0.1877 | 0.1687 | 0.1763 | 0.1901 | 0.2061 |
| \( t^{1/3} \)   | 0.3173 | 0.2216 | 0.2280 | 0.2433 | 0.2588 | 0.2723 |
| \( t \)         | 0.3193 | 0.2155 | 0.1994 | 0.2059 | 0.2179 | 0.2318 |
| \( t^{1/3} \)   | 0.3226 | 0.2209 | 0.1930 | 0.1924 | 0.2012 | 0.2126 |

5. Experimentation

5.1 Creating a Domain Set for Evaluation

The following three factors of a negotiation domain influence the accuracy of preference estimation [7, 8]:

- **Domain size**: Total number of possible bids in the domain. The larger the domain size, the higher the computational cost the model incurs.
- **Bid distribution**: Average distance to the nearest Pareto optimal bid. A large number of bids in a domain with a high bid distribution are far from the Pareto frontier.
- **Opposition**: Distance from the point of best utilities (1.0, 1.0) to the Kalai–Smorodinsky point. The number
HOSOKAWA and FUJITA: OPPONENT’S PREFERENCE ESTIMATION CONSIDERING THEIR OFFER TRANSITION IN MULTI-ISSUE CLOSED NEGOTIATIONS

Fig. 1 Examples of domains used in the experiments. The annotations below show the factors of (Domain size, Bid distribution, Opposition). The yellow points represent Kalai–Smorodinsky points.

of possible agreements will be small if the domain has high opposition.

We created a balanced set of domains with all factors to strictly compare our proposed model and existing models. We created 27 domains such that the three abovementioned factors are large, medium, and small.

We created a set of evaluation domains as follows. First, we determined the number of issues $n$ and the number of options $K_i$ for each issue $I_i$. We set all $K_i$ to be equal to $n$. For example, each issue in a domain with four issues has four options. The weights of each issue and the evaluation values of each choice are randomly generated using probability distribution. We generated issue weights using the Dirichlet distribution as a probability density function. The evaluation values of each issue are generated with the beta distribution. We adjusted the beta distribution parameters according to the desired domain characteristics. We then selected one of the options for each issue and changed the evaluation value to 1 to such that the maximum value in each issue was 1.0.

We randomly created multiple domains, as described above, and selected 27 domains that were balanced in the three factors mentioned above. Figure 1 shows examples of the created domains.

5.2 Experiment 1: Evaluation of the Estimation Accuracy

5.2.1 Experimental Setting

This section compares the proposed TDR model with existing frequency-based models in terms of accuracy. We evaluated the quality of the preference estimation, excluding the influence of both agents’ concession and acceptance strategies in this experiment. As the analysis of ANAC agents in Sect. 4.3 revealed, most of the agents commonly conceded as time passes with using time-based concession functions. Moreover, we could see that the shape of the concession function, which is determined by the parameter $e$, was varied by the agents’ strategies. Therefore, we made 10 conceding agents for the opponent, which conceded depending on the current time $t$ as Eq. (5). Figure 2 shows the target utility of the agents over time.

$$U_{\text{target}}(t) = 1 - (1 - U_{\text{min}}) \times t^{1/e}$$

($U_{\text{min}} \in \{1.0, 0.8, 0.6, 0.4\}, e \in \{0.2, 1.0, 5.0\}$)  \hspace{1cm} (5)

These agents select a bid with the closest utility to the target and did not accept any bid in order to obtain the results of the opponent modeling until the end of a negotiation using the same series of offers as the input.

The number of bids received from the opponent in a negotiation also affects the preference estimation accuracy. Therefore, we define the deadline of a negotiation by the number of rounds, i.e., the number of offers exchanged in a negotiation. We set the deadline as 2000, 4000, 6000, 8000, and 10,000 rounds to simulate the difference of the proposal speed of opponents in real-time negotiations.

We inputted a total of 2700 patterns of bid sequences to the HardHeaded agent’s frequency model, distribution-based frequency model, and TDR model and evaluated the estimation accuracy.

We used the following three accuracy measures to evaluate the estimation models.

- Average difference between the evaluation values:
the average difference between the real evaluation value \( \text{eval}(v^j_i) \) and the estimated evaluation value \( \widehat{\text{eval}}(v^j_i) \) for all options in the domain.

- **Average difference between the issue weights:**
  - the average difference between the real issue weight \( w_i \) and the estimated issue weight \( \hat{w}_i \).
- **Pearson correlation of bids:**
  - the Pearson correlation coefficient between the real utility \( U(\vec{b}) \) and the estimated utility \( \hat{U}(\vec{b}) \) for each bid \( \vec{b} \) in the domain.

*Average difference between the evaluation values* and *average difference between the issue weights* measure the accuracy of each component of estimation. In contrast, *Pearson correlation of bids* measures the accuracy of the overall outcome space.

### 5.2.2 Experimental Results

As outlined above, we quantified the accuracy of the opponent models using three measures. The graphs in Fig. 3 show the overall average accuracy over negotiation time.

Our model is more accurate, except at the very beginning of the negotiation, in (a) *Average difference between the evaluation values*. In the first half of negotiation, a big gap is found between our model and the distribution-based frequency model. Using the offer ratio \( P(v^j_i, t) \) and applying the weighting function \( w(t) \) makes our model more accurate and robust against various diverse environments. Figure 5 shows the average difference between the evaluation values when the opponent concedes later in a negotiation \( (\epsilon = 0.2) \). The estimation error of our model, particularly against competitive agents that concede only when the deadline is approaching, is much smaller than that of the existing models. These agents make non-significant concessions in the first half; hence, the frequency of particular options significantly increases. The accuracy of our model is high in the first half because of the strong effect of Laplace smoothing. In contrast, the accuracy of the distribution-based frequency model is improved in the second half of the negotiation. The proposed method cannot make full use of the bid information because the weighting function decreases the importance of bids over time.

The proposed model also performs better when measuring (b) *Average difference between the issue weights*. The estimation error of our model monotonically decreases, whereas the error of the distribution-based frequency model is large in the beginning. The distribution-based frequency model defines the window size by a number of sequential bids; therefore, the relative size in the negotiation timeline differs by the deadline. On the other hand, our model can observe changes in the opponent’s bids uniformly because it defines the window size by the length of the negotiation time.

A slight difference can be observed between the three models in (c) *Pearson correlation of bids*. The three models can estimate the order of the bids to the same extent; however, all models become less accurate as the negotiation time passes because the opponent sends less-preferred bids later in the negotiation, making the estimation more difficult. Making more accurate estimations using bid sequences that change with the opponent’s concessions is a common task for frequency models.

Figure 4 shows the accuracy of the estimated evaluation values for each deadline’s number of rounds. The distribution-based frequency model becomes less accurate as the length of the deadline increases because it directly uses the frequency number for the estimation function. Our proposed model uses the offer ratio for this purpose; thus, it can estimate the opponent’s preference similarly even if the deadline length (or the offering speed of the opponent) is changed.

### 5.3 Experiment 2: Evaluation of the Estimation Performance

#### 5.3.1 Experimental Setting

In this experiment, we evaluated the proposed model using the agreement of a negotiation.

We applied opponent modeling methods to a naive agent strategy. The modeling agents evaluate all bids sent
from the opponent using the estimated utility function, and find the lowest utility value for the opponent. Then, they set the utility value as their own target value, i.e., they concede at the same speed as the opponent does. The modeling agents calculate the opponent’s estimated utility of bids near their own target. Then, they offer a bid that is estimated to provide the highest utility for the opponent. Every time agents obtain a new offer from the opponent, they also calculate their own utility. They accept the offer if the utility is higher than the target utility.

Each modeling agent negotiated with a non-modeling agent. The opponent agents followed the concession functions defined in Eq. (6). Figure 6 shows the 15 patterns of concession functions used in experiment 2.

$$U_{target}(t) = 1 - (1 - U_{min}) \times r^{t/e}$$

$$U_{min} \in \{0.8, 0.7, 0.6, 0.5, 0.4\}, e \in \{0.2, 1.0, 5.0\}$$  

The opponent agents do not estimate preferences, so they select a bid with the closest utility to the target. If the offer’s utility is higher than the target utility, they accept the bid as the same as the modeling agents.

The negotiation deadline is defined by the number of rounds, similar to that in experiment 1. The deadlines are 2000, 4000, 6000, 8000, and 10,000 rounds.

Herein, we applied the HardHeaded agent’s frequency model, distribution-based frequency model, and TDR model to naive agent strategy and ran a total of 40,500 negotiations.

We evaluated the performance of the opponent models using the average utility value earned by the modeling agents.

5.3.2 Experimental Results

Figure 7 presents the results of experiment 2 against different opponent types.

The figure shows that, for all types of opponent’s strategies, our TDR model achieves higher average utility than both HardHeaded agent’s frequency model and distribution-based frequency model. If the estimation accuracy of the opponent’s utility value is low, the concession speed of the opponent is estimated to be high, thereby accelerating its own concession. We find that a higher estimation accuracy using our TDR model leads to higher utility acquisition.

The gap between the TDR model and the distribution-based frequency model is smaller when the opponent concedes later in the negotiation. The weighting function of our model leads to this result. Our model regards initial offers as more important than later ones by weighting. If the opponent concedes only later in the negotiation, the weight for the option with the highest evaluation value becomes extremely large, and an accurate estimation cannot be performed. Therefore, the TDR model misreads the opponent’s concession and the agent cannot earn a high utility value.
In general, the characteristics of bids and the time of agreement are different by the opponent’s concession. For our future work, the weighting function should change according to the concession of the opponent in order to achieve a higher utility value regardless of the opponent’s strategy.

In conclusion, a negotiating agent that contains our TDR model obtains higher average utility in negotiations than that includes existing models. It is because of the fact of our model’s high accuracy and robustness against various opponents’ strategy, which is figured in experiment 1.

6. Related Work

Numerous learning methods for the opponent have been presented in automated negotiation, including learning the opponent’s offering strategy, acceptance policy, and preferences. Among them, the opponent’s preference modeling is considered to be one of the most important factors for negotiating strategy. In most cases, the opponent model is constructed to reduce the negotiation time, achieve a win–win agreement, or adjust the agent’s strategy toward the opponent. Some researchers recently claimed that a negotiating strategy can be divided into the following parts: offering strategy, opponent modeling, and acceptance strategy [14]. Although an agent’s purpose in estimating the preferences is different, it is suggested that the estimation methods can be compared as detached modules.

Bayesian learning is one approach to estimate the opponent’s preferences. Hindriks and Tykhonov [15] applied the Bayesian learning technique to estimate the shape of the opponent’s utility function (i.e., downhill, uphill, and triangular). That model estimates the issue weights and the evaluation values separately. The opponent’s utility function is presumed to be represented by a combination of simple function shapes. Bayesian learning is used to predict the probability distribution of each utility function in the hypothesis space. Williams et al. [16] also implemented a Bayesian approach to the negotiating agent, IAMHaggler. In that model, the opponent is assumed to concede steadily as time passes. Williams et al.’s model exclusively uses unique bids to update its estimation. The model needs the information about the utility of the opponent’s bids; however the Bayesian learning method is applied. Most models assume the opponent to follow a specific concession function; however, some opponents do not fit such an assumption.

Frequency models are also applied to many agent strategies. These models are based on how often the opponent chooses the options. The most popular model among the negotiating agents is HardHeaded agent’s frequency model [9]. This model estimates the opponent’s evaluation value by counting the frequency of bid selections. The estimation of the issue weights is built with the frequency of changes in the opponent’s offers. Tunahl et al. [10] extended the HardHeaded agent’s frequency model to become more complex. The model separates the bid history of the opponent into consecutive windows and compares the selection distributions between two windows to issue weight estimation. The two abovementioned models were also discussed in detail in Sect. 3. van Galen Last [17] proposed a frequency model similar to the HardHeaded agent’s frequency model. This proposed model uses the frequency of offers to estimate the issue weights and evaluation values. The estimation of the issue weights depends on the ratio of the most selected options. Ikraishi and Fujita [18] focused on multi-time negotiations. Their model estimates the opponent’s preferences using the bid history of past negotiations. Their model considers the time each bid was offered to make an estimation. Frequency models also have to make some assumptions about the opponent’s strategy, but these assumptions are generally weaker than those of Bayesian models.

Other learning techniques are used for opponent modeling in some cases. Zafari and Nassiri-Mofakham [19] proposed perceptron-based opponent modeling, wherein the model considers the opponent’s preferences as a perceptron unit. Each time the agent obtains a new offer from the opponent, the model updates the estimation using a gradient descent search algorithm.

7. Conclusions

Agents keep their preference information private in automated negotiation. Hence, an opponent’s preference model must be created to make a better agreement. Frequency opponent models are widely used for estimation; however, these models are often weak against some opponents. This study proposed a new model for opponent modeling, the TDR model. The main features of the proposed model are the offer ratio and weighting function. Experimental results demonstrated that the proposed model is more accurate than existing frequency models. In addition, the proposed model is more robust against changes in the opponent’s offering speed. We confirmed that the agent applying our model outperformed the other agents in terms of achieving a higher utility during negotiations.

The experimental results clearly showed that the overall accuracy of the frequency models including our model
declined over time because the opponent offers less preferred bids as the opponent concedes. Analyzing the opponent’s concession and constructing the estimation mechanism considering such changes are needed; nonetheless, the novel model we propose does increase accuracy over other existing models. Another direction for future work is to extend our model for multi-time negotiations. Our model uses offer ratio instead of frequency counts to estimate the opponent’s preference similarly even if the offering speed is changed. In multi-time negotiations, we can take advantage of this feature to estimate the opponent’s utility function even if the count of past offers increases due to the accumulation of negotiation history. We will test how our estimation model works when it can refer to the history of past negotiations.

References

[1] S. Fatima, S. Kraus, and M. Wooldridge, Principles of automated negotiation, Cambridge University Press, 2014.
[2] N.R. Jennings, P. Faratin, A.R. Lomuscio, S. Parsons, M.J. Wooldridge, and C. Sierra, “Automated negotiation: prospects, methods and challenges,” Group Decision and Negotiation, vol.10, no.2, pp.199–215, 2001.
[3] R. Aydogan, T. Baarslag, K. Fujita, C. Jonker, and J. Mell, “The tenth international automated negotiating agents competition (anac2019),” http://web.tuat.ac.jp/~katfuji/ANAC2019/, 2019.
[4] R. Aydogan, T. Baarslag, K. Fujita, T. Ito, C. Jonker, D. de Jonge, and J. Mell, “The ninth international automated negotiating agents competition (anac2018),” http://web.tuat.ac.jp/~katfuji/ANAC2018/, 2018.
[5] T. Ito, C. Jonker, R. Aydogan, T. Baarslag, K. Fujita, S. Morinaga, S. Nakadai, T. Yoshida, and Y. Mohammad, “Pacific Rim International Automated Negotiation Agents Competition (PRIANAC),” http://web.tuat.ac.jp/~katfuji/PRIANAC2018/, 2018.
[6] T. Baarslag, M.J. Hendriks, K.V. Hindriks, and C.M. Jonker, “Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques,” Autonomous Agents and Multi-Agent Systems, vol.30, no.5, pp.849–898, 2016.
[7] T. Baarslag, M. Hendriks, K. Hindriks, and C. Jonker, “Measuring the performance of online opponent models in automated bilateral negotiation,” Australasian Joint Conference on Artifi., Intell., pp.1–14, Springer, 2012.
[8] T. Baarslag, M. Hendriks, K. Hindriks, and C. Jonker, “Predicting the performance of opponent models in automated negotiation,” Proc. 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligient Agent Technologies (IAT)-Volume 02, pp.59–66, IEEE Computer Society, 2013.
[9] T. van Krimpen, D. Looije, and S. Hajizadeh, “Hardheaded,” Complex Automated Negotiations: Theories, Models, and Software Competitions, pp.223–227, Springer, 2013.
[10] O. Tunali, R. Aydogan, and V. Sanchez-Anguix, “Rethinking frequency opponent modeling in automated negotiation,” Int. Conf. Principles and Practice of Multi-Agent Systems, pp.263–279, Springer, 2017.
[11] T. Baarslag, R. Aydogan, K.V. Hindriks, K. Fujita, T. Ito, and C.M. Jonker, “The automated negotiating agents competition, 2010–2015,” AI Magazine, vol.36, no.4, pp.115–118, 2015.
[12] R. Aydogan, T. Baarslag, K. Fujita, K. Hindriks, T. Ito, and C. Jonker, “The seventh international automated negotiating agents competition (anac2016),” http://web.tuat.ac.jp/~katfuji/ANAC2016/, 2016.
[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, vol.30, no.1, pp.48–70, Feb. 2014.
[14] T. Baarslag, K. Hindriks, M. Hendriks, A. Dirkzwager, and C. Jonker, “Decoupling negotiating agents to explore the space of negotiation strategies,” Novel Insights in Agent-based Complex Automated Negotiation, pp.61–83, Springer, 2014.
[15] K. Hindriks and D. Tykhonov, “Opponent modelling in automated multi-issue negotiation using bayesian learning,” Proc. 7th international joint conference on Autonomous agents and multiagent systems-Volume 1, pp.331–338, International Foundation for Autonomous Agents and MultiAgent Systems, 2008.
[16] C.R. Williams, V. Robu, E.H. Gerding, and N.R. Jennings, “Lamhaggl er: A negotiation agent for complex environments,” New Trends in Agent-based Complex Automated Negotiations, pp.151–158, Springer, 2012.
[17] N. van Galen Last, “Agent smith: Opponent model estimation in bilateral multi-issue negotiation,” New Trends in Agent-based Complex Automated Negotiations, pp.167–174, Springer, 2012.
[18] M. Ikrashi and K. Fujita, “Compromising strategy using weighted counting in multi-times negotiations,” Advanced Applied Informatics (IIAIAAI), 2014 IIAI 3rd Int. Conf., pp.453–458, IEEE, 2014.
[19] F. Zafari and F. Nassiri-Mofakham, “Popponent: Highly accurate, individually and socially efficient opponent preference model in bilateral multi issue negotiations,” Artifi. Intelli., vol.237, pp.59–91, Aug. 2016.