Adaptive negotiation with on-line prediction of opponent behaviour in agent-based negotiations

Jakub Brzostowski and Ryszard Kowalczyk
Faculty of Information and Communication Technologies
Swinburne University of Technology
Hawthorn, Victoria 3122, Australia
{jbrzostowski, rkowalczyk}@it.swin.edu.au

Abstract

We propose an adaptive approach in agent-based negotiation involving on-line prediction of the opponent behaviour based on the parametric non-linear regression analysis. The predictive decision-making mechanism for the negotiation agent is based on the history of offers in the current negotiation encounter. In comparison to the related work the proposed approach allows the negotiation agents to predict more complex behaviour of the negotiation opponent in terms of mixture of its time-dependant and behaviour-dependant tactics. We perform experiments in order to validate the proposed approach. The results show that the predictive decision-making gives better results in terms of the utility gains for the adaptive negotiation agent as compared with a range of non-predictive negotiation strategies.

1 Introduction

Negotiation is the process of exchanging information among interested parties in order to find an agreement satisfying requirements of these parties [6]. Different fields such as decision and game theory, management and social sciences, artificial intelligence and agent technology [8] study the problem of negotiation from different perspectives. The negotiation agents face various problems such as: limited and uncertain knowledge and conflicting preferences. In addition, in the case of positional bargaining used commonly in agent-based negotiations the agents may have not consistent deadlines and partial overlaps of zones of acceptance. These factors may make reaching an agreement very difficult and there is a need to develop new decision-making mechanisms that can overcome such problems. The agent has to learn and adapt to the behaviour of its negotiation opponent in order to be successful. In the approach proposed by Faratin [5] the agents are equipped in a decision-making mechanism allowing for limited adaptation to the behaviour of negotiation partner. The adaptation is realized by the use of so called behaviour-dependant tactic that imitates the partners behaviour. This decision-making mechanism [5] allows agents to mix time-dependant tactics with the behaviour-dependant tactics using weights what can result in quite complex negotiation behaviour. However, such a negotiation strategy has a large number of parameters that have to be set up by the user and sometimes it is difficult to decide how to set these parameters. Therefore, some approaches based on learning from previous interactions have been proposed.

Most of the work devoted to learning approaches supporting the negotiation focuses on learning from previous encounters, i.e. off-line learning. The approaches include: Bayesian learning [12], Q-learning [4], case-based-reasoning [2] and evolutionary computation [11][10]. However, these approaches require usually history of repeated interactions and such data may not always be available. Therefore, it is important to learn from the history of offers of the current encounter and adapt to the negotiation opponent behaviour, i.e. on-line learning. The online learning approaches supporting the negotiation are quite rare. Hou [7] proposes such an approach based on the non-linear regression analysis. In that work the history of partners concessions is used to predict its future offers and deciding which negotiation strategy to use in order to adapt to the forecast. However, the predicting mechanism is applied only for pure tactics, i.e. either time-dependant or behaviour-dependant.

In this paper we propose an adaptive mechanism based on the regression analysis for predicting more complex behaviour, namely a mixture of the time-dependant and behaviour-dependant tactics. First we forecast the future offers of the negotiation partner as a potential response to our future offers. Having the forecast we then determine a
sequence of our future offers to maximize the expected utility gains. Compared to the approach in our previous work based on the difference method [3] the regression based prediction is more precise and results in higher utility gains of the adaptive negotiation agent. The paper is organised as follows. In the second section we present the possible models of agent behaviour. In the third section we show how the regression analysis is used for the forecasting of the concession curve. In the fourth section we present the way of determining the concession of the adaptive agent based on the forecast from previous section. The fifth section presents the results of experiments and finally we present conclusions and future work.

2 The models of negotiation opponent behaviour

In order to forecast the behaviour of the negotiation opponent a finite set of models is assumed for the opponent agent. Then the time series are fitted to all the models by using regression and the model minimizing the error of the fit is chosen. In this paper we consider four alternative models on the side of the negotiation opponent. We assume that the agent is using the negotiation strategy consisting of time-dependant tactics and behaviour-dependant tactics combined with weights. Therefore we will call this type of strategy two-tactic static strategy. Assuming that a negotiation thread is described as follows:

$$\hat{x}'_{a-b} = f(min^b, max^b, k, t_{max}, \beta, w_t)$$

The value $x_{i-1}$ is an offer proposed by the agent $a$ to agent $b$ in time $t_i$. We will use a simplified notation of the negotiation offers as shown in Table 1

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_1'$</th>
<th>$x_2$</th>
<th>$x_2'$</th>
<th>...</th>
<th>$x_{n-1}$</th>
<th>$x_{n-1}'$</th>
<th>$x_n$</th>
<th>$x_n'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t_1</td>
<td>t_1</td>
<td>t_2</td>
<td>t_2</td>
<td>...</td>
<td>t_{n-1}</td>
<td>t_{n-1}</td>
<td>t_n</td>
<td>t_n</td>
</tr>
<tr>
<td>y_1</td>
<td>y_1</td>
<td>y_2</td>
<td>y_2</td>
<td>...</td>
<td>y_{n-1}</td>
<td>y_{n-1}</td>
<td>y_n</td>
<td>y_n</td>
</tr>
</tbody>
</table>

Table 1. The negotiation thread

A general form of the predictor for forecasting future offers of the negotiation opponent is as follows:

$$f : (H_{t_m}, x_{b-a}^{t_m+1}) \rightarrow x_{a-b}^{t_m+2}$$

where $H_{t_m}$ is the history of behaviour of both agents up to the round $m$ (the sequences of offers of both negotiation parties). The predictor $f$ maps the history and our potential future offer into prediction of negotiation partners offer.

The specific four alternative models of opponents behaviour are defined in the following way:

$$\hat{x}'_{a-b} = f(min^b, max^b, k, t_{max}, \beta, w_t); \quad (1)$$

$$\hat{x}'_{a-b} = f(min^b, max^b, k, t_{max}, \beta, w_t); \quad (2)$$

$$\hat{x}'_{a-b} = f(min^b, max^b, k, t_{max}, \beta, w_t); \quad (3)$$

$$\hat{x}'_{a-b} = f(min^b, max^b, k, t_{max}, \beta, w_t); \quad (4)$$

where the functions $df_{pol}$ and $df_{exp}$ mapping time into $[0, 1]$ are the polynomial and exponential decision functions that are defined in the following way:

$$df_{pol}(t) = \left(\frac{t}{t_{max}}\right)^\beta$$

$$df_{exp}(t) = 1 - e^{-\left(\frac{t}{t_{max}}\right)\beta - 1}$$

The parameters estimated in the model are: $min^b, max^b, k, t_{max}, \beta, w_t$. The parameters $min^b$ and $max^b$ are the boundaries of the zone of acceptance of agent $b$, the $t_{max}$ is the deadline of agent $b$, $\beta$ is the description of the time-dependant decision function (the level of cooperativeness or competitiveness of the agent $b$), $w_t$ is the weight corresponding to the time-dependant tactic and $w_b$ is the weight corresponding to the behaviour-dependant tactic. We can reduce the number of parameters from 6 to 5 (by aggregating the parameter $k$ with the another parameters). The first of the models consists of the mixture of the time-dependant
tactic generated by a polynomial decision function and the absolute tit for tat (ABTFT) tactic of the first order without the random perturbation. The second model is the mixture of the time-dependant tactic generated by an exponential decision function and the ABTFT tactic of the first order without random perturbation. Third model is the mixture of the time-dependant tactic generated by the polynomial decision function and the relative tit for tat (RTFT) tactic of the first order. Fourth model is the mixture of the time-dependant tactic generated by the exponential decision function and the RTFT tactic of the first order. These models are used as predictors of the negotiation opponent behaviour. The data (sequences of offers of both parties) are fitted to all four models and the model minimizing the error of the fit is chosen to forecast the future offers of the negotiation opponent as the responses to our future potential offers.

3 The regression approach for the concession curve forecasting

Assuming the predictor of a general following form:

\[
\hat{z}_{i+2} = f(min^b, max^b, k, t_{max}, \beta, w_t; t_{i+2}, x_{a-b}, x_{b-a}) = f(min^b, max^b, k, t_{max}, \beta, w_t; t_{i+2}, z_i, y_{i+1}, y_{i-1})
\]

we estimate the parameters corresponding to the description of negotiation opponent strategy (\(\beta, w_t\)) and environment (\(min^b, max^b, k, t_{max}\)) using the regression analysis. To simplify the notation we substitute the description of our offers by \(y_t\) and the opponent responses by \(z_{i+1}\) respectively.

The input data of the predictor consists of three time series, namely the series of past time-points \(\{t_i\}_{i \in N}\), the series of our past offers \(\{y_i\}_{i \in 2N+1}\) and the series of the opponent’s past offers \(\{z_i\}_{i \in 2N}\). The output data are the future responses of the negotiation opponent. Before applying the regression we need to transform the problem because the iterative part of the agent behaviour is generated recursively. This means the agent uses its previous offer and previous offers of the opponent to obtain its next proposal. We transform the input data consisting of three time series: \(\{t_i\}_{i \in N}, \{t_i\}_{i \in N}, \{z_i\}_{i \in 2N+1}\) into input data consisting only of two time series \(\{t_i\}_{i \in N}, \{v_i\}_{i \in 2N}\) in the following way:

\[
v_i = z_{i-1} + y_{i-2} + y_t
\]

in the case of model with ATFT

\[
v_i = z_{i-1} \frac{y_{i-2}}{y_t}
\]

in the case of model with RTFT

After transforming the times series the redefined predictor of the agents behaviour has the following form:

\[
z_{i+2} = f(min^b, max^b, k, t_{max}, \beta, w_t; t_{i+2}, v_{i+1}) \tag{6}
\]

\[
= (min:max[w_t[min^b + (max^b - min^b) \times (k df(t_{i+2}) + (1 - k)) + w_b v_{i+1}, min^b], max^b])
\]

The transformed predictor allows us to define the model needed for the regression mechanism in the following form:

\[
z = f(min^b, max^b, k, t_{max}, \beta, w_t; t, v) = \tag{7}
\]

\[
= (min:max[w_t[min^b + (max^b - min^b) \times (k df(t) + (1 - k)) + w_b v], min^b], max^b])
\]

The input variables are \(t\) and \(v\) and the output variable is \(z\). The task of regression analysis is to minimize the error function (the deviation of data from the assumed model) of the following form:

\[
err(min^b, max^b, k, t_{max}, \beta, w_t) = \tag{8}
\]

\[
= \sum_{i \in 2N; i \leq m} (z_{i+2} - f(min^b, max^b, k, t_{max}, \beta, w_t; t_{i+2}, v_{i+1}))^2
\]

The minimization of the error function in the case of the non-linear model requires the application of optimization methods that converge to the optimal solution by the iterative process of correcting the solution (the sequence of estimated parameters). The methods include: the steepest descent [1], the linearisation (Gauss-Newton method) [1] and the Marquardt compromise [9]. The third method [9] is a combination of the first two and the experiments show that it gives the best results in a case of our models.

4 Concession determination

The predictor obtained through regression constitutes a transition function that maps the state of the negotiation (consisting of the history of previous offers of both parties) and an action (which is our potential offer) into the prediction about the next offer of our partner. The predictor can be applied multiple times to asses every sequence of our future actions (potential offers): \(y_{m+1}, y_{m+3}, \ldots, y_{m+l}\). This assessment gives us the sequence of future responses of the negotiation partner: \(z_{m+1}, z_{m+3}, \ldots, z_{m+l}\). Now we need to determine the sequence of offers that will maximize our outcome in terms of the gained utility. We assume that the agents should reach agreement in the last possible round \(t_{max}\) calculated in the following way: \(t_{max} = min(t_{max}^a, t_{max}^b)\) where \(t_{max}^a\) is the deadline of adaptive agent and \(t_{max}^b\) is the prediction about deadline of the non-predictive negotiation partner. Reaching agreement earlier may result in a lower utility because in shorter time the overall concession is usually smaller than the overall concession
made in longer time. Negotiating longer than $t_{max}$ may result in broken negotiation and 0 utility.

The application of “brute force” for the search of optimal sequence of concessions (consideration of all sequences in discrete space) is not feasible because it is too expensive in terms of the computational time. Therefore, more efficient method should be applied. We simplify the search by constraining the sequence of potential offers in the following way. We assume that the first concession may vary and all the concessions from the second one up to the last one are the same. $y_{m+1}, y_{m+1} + c, y_{m+1} + 2c, \ldots, y_{m+1} + (l - 1)c$. Assuming that the first concession was chosen we need to determine the constant $c$ in order to obtain the whole sequence. We do it by solving the optimization problem:

$$c : \min \{ |\bar{z}_{m+l+1} - (y_{m+1} + (l - 1)c)| \}$$

(9)

Various sequences of offers may lead to similar results in the terms of final utility gained by the adaptive agent. Therefore we need to consider all the sequences of offers $\{y^j\}_{i>m, i \in \mathbb{N}+1}$ satisfying the condition 9 (concessions in discrete space).

$$y^1_{m+1}, y^2_{m+1}, \ldots, y^j_{m+l}, \ldots, y^n_{m+1}, \ldots,$$

We can asses every sequence by applying the predictor multiple times and predict the utility that each of the sequences will lead to. Now based on the set of optimal sequences with various first concessions we need to determine the concession for the next round of negotiation. The next concession for each sequence will be determined based on two values: the first concession in the sequence $y^j_{m+2} - y^j_{m+1}$ and the average concession of the sequence $(y^j_{m+i} - y^j_{m+1})/l$.

The more time-dependant (and less behaviour-dependant) the negotiation partner is the less important the first concession is. This means that if our partner is not responding to our behaviour it does not matter what concessions do we offer in what rounds. In the case of high time-dependency the average concession should be taken into account and in the case of high behaviour-dependency the first concession should be considered. The regression mechanism gives us the levels of time-dependency and behaviour-dependency in the behaviour of the non-predictive negotiation partner. These weights can be used as proportions in the aggregation of the two types of concessions for each sequence $\{y^j\}_{i=m, i \in \mathbb{N}+1}$ in the following way:

$$offer^j = y^j_{m-1} + \frac{y^j_{m+l} - y^j_{m-1}}{l} \times w_l + (y^j_{m+1} - y^j_{m-1}) \times w_b$$

The offers $offer^j$ are then averaged over all the optimal sequences with the various first concessions in order to calculate the final value of an offer:

$$offer = offer^1 w_1 + offer^2 w_2 + \ldots + offer^n w_n$$

where the weights $w_j$ are proportional to the utility that can be gained by applying $j$-th sequence of actions $\{y^j\}$ according to our prediction.

5 Results

The mechanism has been tested in a simulated experimental negotiations with more advanced settings than in [5]. On a side of our negotiation agent the predictive mechanism is used and the negotiation opponent uses the non-predictive strategies being a mixture of a time-dependant tactic and a behaviour-dependant tactic with various weights assigned to these two tactics. In order to make experiments feasible we choose seven time-dependant tactics and divide them into three subsets:

- Conceder: $C = \{ \beta \mid \beta \in \{2, 5, 8\} \}$
- Linear: $L = \{ \beta \mid \beta \in \{1\} \}$
- Boulware: $B = \{ \beta \mid \beta \in \{0, 0.5, 0.7\} \}$

We consider two behaviour-dependant tactics:

- Absolute Tit-For-Tat: $a: \delta \in \{1\}$ and $R(M) = 0$ (without a random factor)
- Relative Tit-For-Tat: $r: \delta \in \{1\}$

The space of weights assigned to the tactics consists of 9 values obtained through discretization. Analogously as for the set of $\beta$ values we decompose the set of all weights (corresponding to the time-dependent tactic) into three subsets:

- Small: $S = \{0.1, 0.2, 0.3\}$
- Medium: $M = \{0.4, 0.5, 0.6\}$
- Large: $L = \{0.7, 0.8, 0.9\}$

The cartesian product of the sets of the time-dependant tactics, the behaviour-dependant tactics and the weights results in a set of 126 static strategies:

$$ST = (C \cup L \cup B) \times \{a, r\} \times (S \cup M \cup L)$$

The whole set of the static strategies $ST$ is divided into 18 groups of strategies in order to illustrate the performance of our approach:

$$ST = (C \times \{a\} \times S) \cup (L \times \{a\} \times S) \cup (B \times \{a\} \times S) \cup (C \times \{a\} \times M) \cup (L \times \{a\} \times M) \cup (B \times \{a\} \times M) \cup (C \times \{a\} \times L) \cup (L \times \{a\} \times L) \cup (B \times \{a\} \times L) \cup (C \times \{r\} \times S) \cup (L \times \{r\} \times S) \cup (B \times \{r\} \times S) \cup (C \times \{r\} \times M) \cup (L \times \{r\} \times M) \cup (B \times \{r\} \times M) \cup (C \times \{r\} \times L) \cup (L \times \{r\} \times L) \cup (B \times \{r\} \times L)$$

Each group will be shortly denoted using only the first letters. For example:

\[ \text{CSA} = C \times \{ a \} \times S \]

We consider the scenarios with following values of the environmental variables.

- the client-agent: \( \min^b \in BL = \{10\}, \max^b \in BU = \{50\}, t_{\text{max}}^b \in BD = \{15, 20, 25, 30\} \)

- the provider-agent: \( \min^s \in SL = \{10\} = \{10, 23.2, 36.4\}, \max^b \in SU = \{50\}, t_{\text{max}}^s = t_{\text{max}}^b \)

The results for other scenarios will be presented in another paper. The performance of our modelling approach is illustrated in Figures (1) (2) (3) (4). The blue bars in the top chart represent an average value of utility of a negotiation agent using all the strategies from particular subset of static strategies (for instance: CSA) against all the strategies from the set \( ST \). This corresponds to playing with a strategy from a particular subset against a random strategy selected from the whole \( ST \) according to an uniform probability distribution. The blue bars in the bottom chart represent the maximal value of utility of a negotiation agent using all the strategies from particular subset of static strategies against all the strategies from the set \( ST \). This corresponds to playing the strategy from a particular subset against the best strategy selected from the whole set \( ST \). The red bars in the charts illustrate the utility gain of our adaptive mechanism played against the strategies from a particular subset of strategies.

The strategies with high contribution of time-dependant part (\( BLA, LLA, CLA, BLR \ldots \)) (Figure (1)(2)(3)(4)) are lowly responsive what means that larger concessions on the side of adaptive mechanism will not encourage the negotiation partner to approach our position faster. The adaptive mechanism uses this information and makes small concessions during the whole encounter because it knows that the partner will concede later in the negotiation (Figure (5)). In the case of full overlap of zones of acceptance and consistent deadlines such a behaviour gives very good results (Figure (1)(2)(3)(4)). It can be noted that the utility gained by the adaptive mechanism significantly exceeds the utility gain of the random static strategy and is close to the utility gained by the best static strategy. In the case of other scenarios (not consistent deadlines, partial overlap of zones of acceptance) the problem is more difficult because regression does not predict the reservation value and deadline in the case of polynomial decision function.

The higher the behaviour-dependency of the negotiation partner the more the behaviour of adaptive mechanism differs. In the case of strategies with medium (\( BMA, LMA, CMA, BMR, \ldots \)) and high (\( BSA, LSA, CSA, BSR \ldots \))
contribution of behavioural part (Figure (1)(2)(3)(4)) the concessions of the adaptive mechanism are higher than in the case of strategies with low behaviour-dependency (Figure (5)). The adaptive mechanism takes advantage of the higher behaviour-dependency by making significant concessions in the beginning of the encounter. This encourages the partner to respond positively by making larger concessions. Such a behaviour results in fast agreements with good utility gain (Figure (5)). The utility gain in comparison to the random strategy gain is significantly larger (Figure (1)(2)(3)(4)) and in comparison to the best static strategy it is slightly lower.

The static strategies with the relative tit for tat differ from the strategies with absolute tit for tat. The strategies with ATFT applied on the side of client and provider work in the same way (are symmetrical). However, the RTFT results in different behaviour when applied for client and when applied for provider. It can be noted that when the provider agent is using a strategy LSR (Figure (6)) it responds with high concessions in relation to the concessions of our adaptive mechanism. In the case of LSR applied on the side of client-agent (Figure (7)) the negotiation partner (client) is still responding positively but its concessions are lower.
in the comparison to the concessions of the adaptive mechanism. In the case of both roles profiles the adaptation to highly imitative agent results in quick agreements. However, the resulting utility gains are not symmetrical (Figure (1)(2)(3)(4)) because the client using adaptive mechanism gains $0.625$ of utility and the provider using adaptive mechanism gains $0.59$ of utility.

6 Conclusions and future work

We proposed an approach for the prediction of complex behaviour of negotiation agent based only on the history of offers of current encounter. The prediction mechanism is based on non-linear regression analysis. We also proposed a decision-making mechanism using the information obtained by regression mechanism. We tested the predictive decision-making approach against static two-tactic strategies. The presented results of the experimentation show that the adaptive mechanism outperforms the random static strategy by achieving substantial gains in terms of utility. In the future work we will test the approach in different scenarios involving not consistent deadlines and partial overlaps of zones of acceptance of the negotiation parties.

References