POPPONENT: Highly Accurate, Individually and Socially Efficient Opponent Preference Model in Bilateral Multi Issue Negotiations (Extended Abstract)*

Farhad Zafari and Faria Nassiri-Mofakham*
Swinburne University of Technology, Melbourne, Australia
*University of Isfahan, Isfahan, Iran
fzafari@swin.edu.au, fnasiri@eng.ui.ac.ir

Abstract
In automated bilateral multi issue negotiations, two intelligent automated agents negotiate on behalf of their owners over many issues in order to reach an agreement. Modeling the opponent can excessively boost the performance of the agents and increase the quality of the negotiation outcome. State of the art models accomplish this by considering some assumptions about the opponent which restricts their applicability in real scenarios. In this paper, a less restricted technique (POPPONENT) is proposed, where perceptron units are applied in modelling the preferences of the opponent. This model adopts a Multi Bipartite version of the Standard Gradient Descent search algorithm (MBGD) to find the best hypothesis, which is the best preference profile. In order to evaluate the accuracy and performance of this proposed opponent model, it is compared with the state of the art models available in the Genius repository and in the devised setting. The results approve the higher accuracy of POPPONENT compared to the most accurate state of the art model. Evaluating the model in the real world negotiation scenarios in the Genius framework also confirms its high accuracy in relation to the state of the art models in estimating the utility of offers. The findings here indicate that the proposed model is individually and socially efficient. The proposed MBGD method could also be adopted in similar practical areas of Artificial Intelligence.

1 Introduction
Negotiation is the science and art of resolving any kind of disputes and reaching consensus among human parties. In an automated bilateral multi issue version of negotiations, intelligent computer agents engage in a cooperative process on behalf of their beneficiaries with different and sometimes contradicting interests, with the objective of achieving an agreement on one or more issues. Recently, with the emergence of ANAC (an annual international Automated Negotiating Agents Competition) [Baarslag et al., 2012; Fujita et al., 2013], many new negotiation strategies have been developed. Most of the existing sophisticated negotiation strategies typically consist of a set of fixed modules. In general, three main components are distinguished in a negotiating agent which work together within a BOA framework to accomplish the whole negotiation task [Baarslag et al., 2014]: Bidding Strategy (BS), which decides which offer will be sent to the opponent as the next proposal, Opponent Model (OM), which constructs a model of the preference profile of the opponent through a learning technique, and Acceptance Strategy (AC), which receives the incoming offer from the opponent and the outgoing offer chosen by the bidding strategy component, and determines whether the incoming offer is acceptable for the agent or not.

Despite the variety in opponent modelling techniques, most of the current models rely on a small and common set of learning techniques [Baarslag et al., 2016]. Moreover, since in a single negotiation session all bids (i.e., training examples) are not available at the same time, traditional learning techniques are not easily applied. An opponent model that is able to learn incrementally and update itself once the new training examples (i.e., offers) arrive during a negotiation session is of necessity. Another problem in the negotiation setting is that the training instances lack the output variable (i.e., the variable which contains the utility values of the received bids in the opponent’s view). Therefore, specific opponent models are required that are capable of modeling the preferences of the opponent with no need for the value of the output variable. In order to overcome this limitation, all the existing opponent models use a subset of assumptions to extract the preferences of the opponent [Zafari and Nassiri-Mofakham, 2016b]. Another difficulty with modeling an opponent’s preferences in bilateral negotiations is related to the time factor. The post event analysis of ANAC tournaments also confirms that the computational complexity of the opponent models and the poor accuracy are the two main factors that degrade the performance of the agents applying these models [Baarslag, 2016]. In particular, the time factor is of paramount importance in online opponent models. In these models, the participating agents usually exchange a limited number of offers before the negotiation deadline is met, therefore, they do not contain enough information to accurately train an opponent model [Hindriks and Tykhonov, 2008;
Consequently, the ability of the model to extract the most information possible from the training bids it receives is highly essential. Therefore, a proper opponent model should function based on the least assumptions, be efficient, extract the most information from minimum number of bids, and embrace incremental training capability. To learn the issue weight values and individual utility function captured by opponent models, a two layered architecture would be essential. To overcome the aforementioned difficulties, in this paper, a new opponent preference model based on perceptron units is proposed in bilateral multi issue negotiation domains. Moreover, to be more applicable in real world negotiations, fewer and more realistic assumptions than those of the state of the art models are applied in this study. This is obtained by proposing an opponent model named POPPONENT, based on an adapted version of the standard gradient decent search [Zafari and Moser, 2016; Zafari and Moser, 2017b; Zafari and Moser, 2017a] (named the Multi Bipartite Gradient Decent Search). The model shows great success in the practical AI area of modelling the opponent preferences in bilateral multi issue negotiations with incomplete information, and the proposed Multi Bipartite Gradient Decent Search method used to train the model can be conveniently adopted in the similar AI problems.

2 Negotiation Setting

The negotiation setting here is in accordance with the setting employed by the state of the art models in the field of bilateral automated negotiations and the setting of the ANAC 2010-2013 [Baarslag et al., 2012; Fujita et al., 2013; Lin et al., 2014; Baarslag et al., 2013; Yaakov and Ilany, 2015; Zafari et al., 2015].1 Automated agents alternatively exchange offers and compete against each other to reach a joint agreement on a set of issues in bilateral negotiations. The issues and possible values for each issue constitute a domain. For each domain there could be two preference profiles (one for each side of the negotiation) which together with the domain construct a negotiation scenario. The interaction between negotiating parties is regulated by a negotiation protocol that defines the rules of how and when proposals can be exchanged. In this setting, the alternating offers protocol is applied [Rubinstein, 1982]. Negotiation is bilateral, that is, exactly two parties are negotiating over one or a set of issues. Each issue is associated with a set of possible values. The agents repeatedly exchange offers in successive rounds, so as to reach a mutually acceptable outcome. The negotiation deadline is reached after a specified number of $N$ rounds are passed. This type of negotiation setting is commonly referred to as a round based setting. Each agent tries to take advantage of the other party for gaining a maximum utility for its own. A negotiation break-off causes both negotiating parties to obtain their reservation values. Therefore, the agents try to reach an agreement before the deadline. A negotiation session takes place in a negotiation scenario, which consists of a negotiation

\[\text{domain (or, alternatively, an outcome space) and two preference profiles (or, alternatively, utility space) one for each negotiating agent. The negotiation domain } \Omega \text{ specifies all possible offers } \vec{o} \text{ that the agents can send or receive. Each offer or possible outcome is a vector } < o_1, ..., o_k >, \text{ where each component is the mapping of every issue } i \text{ to a value } o_i \in [v_{i1}, v_{im}], \text{ where, } m_i \text{ is the number of possible values for issue } i, i = 1, ..., n [\text{Fujita et al., 2013}]. \text{ A preference profile } \{ < \vec{o}, U(\vec{o}) > | \vec{o} \in \Omega \}, \text{ on the other hand, consists of a utility function } U(\vec{o}) \text{ which maps each possible offer } \vec{o} \in \Omega \text{ to a value in the } [0,1] \text{ range based on the overall relative value of that offer for the agent. In multi-issue negotiations, the common assumption is that the utility of an offer can be computed as a weighted sum of the utilities associated with the values for each issue [Nassiri-Mofakham et al., 2008; Nassiri-Mofakham et al., 2009]. Accordingly, in the negotiation setting here, the agents use the linear additive utility function shown in Equation (1), defined by a set of weights } w_i, \text{ and the corresponding evaluation functions or evaluation values } eval_i(o_i), i = 1, ..., n \text{ for the issue value } o_i \text{ of a given offer } \vec{o}. \]

\[U(\vec{o}) = \sum_{i=1}^{n} w_i \times eval_i(o_i) \quad (1)\]

Unlike the negotiation domain which is publicly known for both the negotiating parties, the preference profile is private for each agent, so the agents are not aware of the weights and evaluation values associated with the preference profile of one another. This negotiation setting is online, meaning that the agent is only allowed to use the offers exchanged during a single negotiation session to model the preferences of the opponent. Unlike offline opponent models, where negotiation information from different negotiation sessions is used, in online models [van Krimpen et al., 2013] no history of the previous negotiations is provided for the opponent model.

3 POPPONENT Algorithm

The preference profile of the opponent in linear scenarios is learnt using the multi bipartite incremental gradient descent search [Zafari and Nassiri-Mofakham, 2016b]. This algorithm learns the issue priorities or weight values $w_1^{OP}$ through $w_n^{OP}$ and the evaluation values $eval^{OP}(\omega_i)$ for all possible values $\omega_i \in [v_{i1}, ..., v_{im}]$ (where, $m_i$ is the number of possible values for issue $i$) and all negotiation issues $i (i = 1, ..., n)$ in that negotiation domain.

This proposed algorithm implements the multi bipartite incremental gradient descent search and applies two parameters of $\eta$ and $N$ as the input. Parameter $\eta$ represents the learning rate which determines the step size in the gradient descent search. Parameter $N$ represents the number of training repeats for each training instance. This algorithm includes two separate functions of initializer and updater. The first function is invoked just once when the model is generated and the essential parameters of the proposed model, most importantly the preference profile of the opponent (that is, the issue weights $w_i^{OP}$ and evaluations $eval^{OP}(\omega_i)$ for issue values) is initialized. By trying different initial points, it is realized that 0.5, the midpoint in the hypothesis space, is the best point for initializing $eval^{OP}(\omega_i)$ values for each $\omega_i \in [v_{i1}, ..., v_{im}], i = 1, ..., n$
(Algorithm: line 13). Similarly, for the weight values \( w_i \), equal weights \( \frac{1}{n} \) are chosen for all issues (Algorithm: line 14). The second function receives an offer vector \( \vec{\omega} \) which specifies the issue values for all the negotiation issues of a new offer recently received from the opponent. That is, as soon as a new offer is received from the opponent, this function is invoked to update the model based on this newly received bid. It updates the estimated preference profile of the opponent by adjusting \( \text{eval}^{\text{OP}}(\omega_i) \) and \( w_i^{\text{OP}} \) values (Algorithm: Lines 18 and 21).

Whenever a new offer is received from the opponent, the perceptron learning delta rules (Algorithm: Lines 18 and 21) are repeated \( N \) times. The \( \text{EstimatedUtility}^{\text{OP}}(\omega) \) is a function which receives an offer as the input and returns the estimated utility value of that offer in the opponent’s utility space as the output. In this algorithm, instead of updating each \( \text{eval}^{\text{OP}}(\omega_i) \) and \( w_i^{\text{OP}} \) value after calculating all \( \Delta w_i \) and \( \Delta \text{eval}^{\text{OP}}_i \) values when all training examples are met, each \( \text{eval}^{\text{OP}}(\omega_i) \) and \( w_i^{\text{OP}} \) value is modified using training delta rules right after each single training instance is met in an incremental manner (Algorithm: Lines 18, 21). Therefore, this algorithm can easily be applied in more realistic negotiation scenarios in which training examples (opponent offers) are gradually met one at a time. The POPPONENT algorithm is as follows:

**Algorithm: Perceptron-Based Opponent Model**

1. \( \text{EstimatedUtility}^{\text{OP}}(\omega) \) is the function which deals with the first sub-problem, can be estimated according to the perceived bidding behavior of the opponent. In this article, four different values - three constant (0.6, 0.8, and 1) and one adaptive method (named P6, P8, P1, AP, respectively) are applied in order to estimate the utilities of the offers proposed by the opponent. For the fourth value, we use the adaptive method where the agent estimates the bids that the opponent will offer in future, based on the opponent’s bid history [Ikraši and Fujita, 2014; Witten et al., 2016]. The computational complexity of POPPONENT Algorithm is linear \( (O(n)) \).

4 Experiments

To evaluate the proposed model, two separate experimental settings are applied for assessing its accuracy and performance, in real world negotiation examples, compared to the available opponent models from Genius repository [Lin et al., 2014]. In the following sections, we present the experimental results obtained in these two experimental settings (for the details, please see [Zafari and Nassiri-Mofakham, 2016b]). The models used in these experiments are abbreviated in Table 1.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Opponent Model</th>
<th>Abbreviation</th>
<th>Opponent Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>P6</td>
<td>Perceptron Based Model (Constant 0.6)</td>
<td>SF</td>
<td>Smith Frequency Model</td>
</tr>
<tr>
<td>P8</td>
<td>Perceptron Based Model (Constant 0.8)</td>
<td>FF</td>
<td>The Fokken Frequency Model</td>
</tr>
<tr>
<td>P1</td>
<td>Perceptron Based Model (Constant 1)</td>
<td>IHB</td>
<td>JAI/Haggler Bayesian Model</td>
</tr>
<tr>
<td>AP</td>
<td>Adaptive Perceptron Based Model</td>
<td>PHRB</td>
<td>Prefit IA/Haggler Bayesian Model</td>
</tr>
<tr>
<td>PP</td>
<td>Perfect Perceptron Based Model</td>
<td>NRB</td>
<td>The Negotiator Random Bayesian Model</td>
</tr>
<tr>
<td>LGF</td>
<td>AgentLG Frequency Model</td>
<td>SB</td>
<td>Scaleable Bayesian Model</td>
</tr>
<tr>
<td>XF</td>
<td>AgentX Frequency Model</td>
<td>PSN</td>
<td>Prefit Scalable Bayesian</td>
</tr>
<tr>
<td>CKF</td>
<td>CI2Agent Frequency Model</td>
<td>OM</td>
<td>Opposite Model</td>
</tr>
<tr>
<td>HHP</td>
<td>HandHeaded Frequency Model</td>
<td>NM</td>
<td>No Model</td>
</tr>
<tr>
<td>IFX</td>
<td>Inn-o-Agent Frequency Model</td>
<td>PM</td>
<td>Perfect Model</td>
</tr>
<tr>
<td>NF</td>
<td>Nash Frequency Model</td>
<td>WM</td>
<td>Word Model</td>
</tr>
</tbody>
</table>

Table 1: The list of the opponent models used in the experiments

### 4.1 Experiment I

The first setting evaluates the accuracy of POPPONENT model through the Pearson Correlation measure. The experimental setting applied by [Baarslag et al., 2013] for automated bilateral multi issue negotiations is applied to evaluate the accuracy of POPPONENT model versus the state of the art opponent models. According to this setting, 5 variations of POPPONENT are compared with a total of 15 opponent models. In this experiment, the agents employing these opponent models negotiate against a category of opponent agents with different bidding strategies in a number of negotiation domains. The average accuracies of top performing opponent models against all opponent agents are depicted in Fig. 1, where at the perfect information state (PP), the proposed model outperforms the state of the art models by a large margin. As observed, the other three variations of POPPONENT (i.e., P6, P8, and P1 except PP) outperform the state of the art models with respect to the average accuracy over all opponents.

### 4.2 Experiment II

In the second Experiment, the proposed model is evaluated by measuring the real performance of the agents applying this model in 7 domains (Table 2) of real world experimental
Fig. 1: Average accuracy of POPPONENT variations and the top state of the art opponent model in Experiment I

Table 2: Top performing real model in each domain for all performance/accuracy measures in Experiment II

<table>
<thead>
<tr>
<th>Domain</th>
<th>Utility</th>
<th>Pearson</th>
<th>Nash</th>
<th>Pareto</th>
<th>Kalai</th>
<th>Time</th>
<th>Pareto Bids %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supermarket</td>
<td>IXF P1</td>
<td>IXF</td>
<td>IXF</td>
<td>IXF</td>
<td>IXF</td>
<td>IXF</td>
<td></td>
</tr>
<tr>
<td>Travel</td>
<td>IXF P1</td>
<td>IXF</td>
<td>IXF</td>
<td>IXF</td>
<td>AP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson</td>
<td>IXF P1</td>
<td>CKF</td>
<td>IXF</td>
<td>CKF</td>
<td>CKF</td>
<td>P1</td>
<td></td>
</tr>
<tr>
<td>Grocery</td>
<td>AP</td>
<td>IXF</td>
<td>CKF</td>
<td>AP</td>
<td>SF</td>
<td>LGF</td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>HHF P1</td>
<td>HHF</td>
<td>IXF</td>
<td>IXF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camera</td>
<td>LGF P1</td>
<td>CKF</td>
<td>LCF</td>
<td>LCF</td>
<td>LCF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ItexVsCypress</td>
<td>IXF</td>
<td>AP</td>
<td>IXF</td>
<td>AP</td>
<td>IXF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>IXF</td>
<td>IXF</td>
<td>IXF</td>
<td>IXF</td>
<td>LF</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Average performance of models in terms of a) Utility, b) Nash Distance, c) Kalai Distance, d) Pareto Distance, and e) Pearson Correlation in Experiment II

5 Conclusion

In this paper, a new opponent preference model called POPPONENT is proposed. The experiments also reveal that the accuracy of P1, P6, and P8 exceeds the accuracy of the most accurate state of the art model. Evaluating the performance of POPPONENT also indicated that it overcomes the most accurate state of the art opponent models. In particular, the results indicated that POPPONENT works better in medium to large and more distributed negotiation domains and overcomes all the state of the art models in at least one domain for all the measures.
References


