

Towards a Quality Assessment Method for Learning Preference Profiles in Negotiation

Koen V. Hindriks and Dmytro Tykhonov

EEMCS, Delft University of Technology, Delft, The Netherlands
{k.v.hindriks, d.tykhonov}@tudelft.nl

Abstract. In automated negotiation, information gained about an opponent's preference profile by means of learning techniques may significantly improve an agent's negotiation performance. It therefore is useful to gain a better understanding of the factors that influence the quality of learning. The benefits of learning in negotiation are typically assessed indirectly by means of comparing the utility levels of agreed outcomes and other more global negotiation parameters. An evaluation of learning based on such general criteria, however, does not provide any insight into the influence of various aspects of negotiation on the quality of the learned model itself. The quality may depend on such aspects as the domain of negotiation, the structure of the preference profiles, the negotiation strategies used by the parties, and others. To gain a better understanding of the performance of proposed learning techniques in the context of negotiation and to be able to assess the potential to improve the performance of such techniques a more systematic assessment method is needed. In this paper we propose such a systematic method to analyse the quality of the information gained about opponent preferences by learning. The method includes measures to assess the quality of a learned preference profile and proposes an experimental setup to analyse the influence of various negotiation aspects on the quality of learning. We apply the method to a Bayesian learning approach for learning an opponent's preference profile and discuss our findings.

1 Introduction

In the area of learning in negotiation, the benefits of the application of learning techniques are often measured with respect to the final outcome or other generally relevant parameters of negotiation such as number of rounds to reach an agreement. The quality of the model of what has to be learned thus is often not directly analyzed but a more indirect method is used to assess these benefits. One of the problems with this indirect method of measuring the benefits of learning in negotiation is that it does not provide any tools to analyse the performance of the learning mechanism itself nor does it provide insight into the factors influencing the quality of learning in negotiation.

To gain a better understanding of the performance of proposed learning techniques and the potential to improve the performance of such techniques in the context of automated negotiation a more systematic assessment method for the quality of learning is needed. Such a method should provide the technical tools

for analysis, identify the key factors that need to be taken into account and propose an experimental setup to evaluate the quality of learning. In this paper we present a method that can be used to assess the quality of a learning of an opponent's preference profile. Useful technical tools as well as an approach for analysis are discussed, and we apply the method to illustrate the results and some of the insights that may be gained by using it.

The paper is organized as follows. In Section 2 we briefly discuss learning in negotiation and introduce the problem concerning the quality of learning in negotiation that we address in this paper. In Section 3 a method is proposed and the various components of the proposed method are discussed. In Section 4 the proposed method is applied to the Bayesian learning approach proposed in [8] and results are presented. In Section 5 conclusions and several directions for future research are outlined.

2 Related Work and Problem Description

Learning in automated negotiation is an important topic since it has been shown that it can significantly improve the performance of a negotiating agent. Work in the area of opponent modelling in negotiation has resulted in a variety of approaches that usually focus on learning one aspect of the negotiation process. The range of negotiation aspects that are learned includes reservation values [21], issue priorities (or weights associated with negotiated issues modelling the relative importance of each issue; [2, 9]), and negotiation strategies [12, 14].

In [11] an opponent's preference profile is learned in a qualitative negotiation setting. It is assumed that a fixed set of possible profile types is given. Bayesian learning then is used to determine the likelihood that an opponent has one of these given profiles. In [4] a model is presented that incorporates domain knowledge for deciding on a negotiation move, which is extended in [2] with a learning technique based on kernel density estimation (KDE) to learn the issue priorities of an opponent. [9] proposes an alternative method for learning issue priorities. In [8] a Bayesian learning technique is presented to learn an opponent's preference profile including both issue priorities as well as the ranking of issue alternatives. The evaluation method that has been used to assess the quality of learning in each of these approaches has been indirect, e.g. by evaluating the improvement of the outcome that is reached with respect to standard notions such as Pareto efficiency. An exception is the work reported in [17] where the quality of learning is discussed on the basis of statistical analysis, and the work reported in [1] that presents quantitative results on a Bayesian classifier to classify the type of profile of an opponent. However, it is not clear from [1, 17] how various factors determine the quality of learning.

In order to define a quality assessment method that provides insight into the contribution of various factors to the learning quality, we first introduce the model of negotiation that we use. Negotiation is a form of decision-making where two or more parties jointly search a space of possible outcomes Ω with the goal of reaching a consensus [18]. In this paper, we only consider *bilateral*

negotiation, i.e. negotiation between two parties. We further assume that both parties are able to express their preferences over possible outcomes $\omega \in \Omega$ and that these preferences can be modelled by means of a utility function U that maps a possible outcome ω to a real-valued number in the range $[0; 1]$ (cf. [20]). A utility function will also be referred to as a *preference profile*.

Possible outcomes of a negotiation may have additional structure and consist of a package-deal of several *issues* or attributes. Each issue has an associated range of alternatives one of which for each issue needs to be agreed upon to reach a final outcome. The space of possible outcomes each of which consists of values assigned to a number of issues is also called the *negotiation domain*.

It is often assumed that a preference profile can be defined as a function of the evaluation functions associated with individual issues and we do so as well here. More specifically, we assume that utility functions are linearly additive [16]. That is, in a domain with n issues and outcomes that consist of one alternative x_i for each of the n issues, we assume that a utility function can be defined by:

$$U(\omega) = \sum_{i=1}^n w_i e_i(x_i \in \omega)$$

where the w_i are normalized weights that sum to 1 and the $e_i(x_i \in \omega)$ are evaluation functions with range $[0; 1]$ which model preferences for issue alternatives. An important reason that justifies this restriction is that most existing negotiation strategies can handle linearly additive utility functions but cannot (efficiently) handle more complex utility functions.

In order to obtain an advantageous negotiation outcome, i.e. to reach an agreement as best as possible, it is useful to have as much information about the preference profile of an opponent as is possible. In a closed negotiation the negotiating parties however do not exchange information about the preferences of each other. In single-instance negotiations a negotiating agent may then try to obtain a model of the preference profile of its opponent by means of learning [2, 8, 9, 11]. The goal of applying learning techniques here is to construct a function \tilde{U} that is similar to the actual utility function U of the opponent. The problem that we address in this paper is how to assess the quality of a learning technique in this context, that is, which tools can be used to assess the similarity of the learned preference profile with the actual profile and which factors influence the similarity. The method proposed aims at a direct assessment of the quality of a learned preference profile instead of indirect evaluations based on results that indicate comparative utility increases of negotiation outcomes, reaching agreements in fewer negotiation rounds, or outcomes closer to the Pareto frontier or fair outcomes such as the Nash outcome. One of our objectives is to be able to analyze the influence of various negotiation aspects on the learning quality.

3 Quality Assessment Method

The method we propose has three components: (i) quality measures to estimate the learning performance, (ii) criteria for selecting a diverse range of negotiation

domains and preference profiles on these domains, and (iii) criteria for selecting a number of negotiation strategies of the opponent. These components then are used to define an experimental setup to obtain data to analyze learning quality by means of a negotiation tournament.

The first component consists of several similarity measures that provide a metric for assessing the accuracy of the learned preference profile with respect to the actual preference profile. We discuss several measures that can be used to assess the quality of the learned preference profile. Apart from the restriction on utility functions which need to be linearly additive, the second component of the method consists of several additional criteria for selecting negotiation domains such as size and complexity of the domain, and the similarity of the preference profiles of the negotiating parties. These criteria are used to define the experimental setup of the negotiation tournament. The third component provides criteria for selecting negotiation strategies that should be used by negotiating agents in the tournament. Since learning of an opponent’s preference profile in single-instance negotiations has to be accomplished with only the observations of the opponent’s negotiation moves [8, 12, 17, 22], typically such learning algorithms use assumptions about an opponent’s behaviour. For instance, in [1, 8, 22] a concession assumption is used which states that negotiators on average decrease the utilities of offers as time passes in order to find a deal. Although this assumption is reasonable and can be applied in typical negotiation settings, it is important to assess the robustness of a learning technique also when negotiating against agents that use strategies that do not comply with this assumption. It thus is important to incorporate a diverse range of negotiation strategies in any experimental setup to evaluate learning quality.

3.1 Quality Measures

In this Section we discuss two quality measures to assess learning quality that are based on two metrics to measure the distance between the actual preference profile of an opponent and the learned preference profile. These quality measures are applied to both the complete preference profiles or utility functions, as well as to the issue priorities or weights.

The learning task of learning an opponent’s preference profile clearly is an approximation problem. The task is to re-constructs the actual utility function U of the opponent by means of a learning technique resulting in an approximate function \tilde{U} . A quality measure with respect to learning preference profiles therefore can be defined as a distance metric of two utility functions, and can be formally represented as $d(U, \tilde{U})$.

Ideally, the approximation \tilde{U} of an opponent’s utility function would provide an accurate prediction of the exact utility value an opponent associates with an outcome. Some strategies like the Tit-for-Tat-based strategy introduced in [3] depend on the accuracy of cardinal values of the utility function of the opponent since a negotiation move is chosen based on an estimate of the concession the other party made in the previous move. It therefore is important to have a distance metric that can be used to measure the accuracy of the cardinal values

predicted by the learned profile. Here we use Pearson's correlation coefficient for that purpose. This coefficient represents the degree of linear relationship between two variables and is defined as follows:

$$d_{pearson}(U, \tilde{U}) = \frac{\sum_{\omega \in \Omega} (U(\omega) - \langle U \rangle)(\tilde{U}(\omega) - \langle \tilde{U} \rangle)}{\sqrt{\sum_{\omega \in \Omega} (U(\omega) - \langle U \rangle)^2 \sum_{\omega \in \Omega} (\tilde{U}(\omega) - \langle \tilde{U} \rangle)^2}} \quad (1)$$

where $\langle U \rangle$ (respectively $\langle \tilde{U} \rangle$) denotes the average utility over the outcome space defined by utility function U (\tilde{U}). The Pearson's correlation coefficient takes a real value from the interval $[-1; 1]$. A value of $+1$ means that there is a perfect positive linear relationship between variables, whereas a value of -1 means that there is a perfect negative linear relationship between variables. A value of 0 means that there is no linear relationship between the two variables.

Although a perfect match of cardinal values of the actual and learned utility function would be ideal, in practice it may be sufficient and more important to approximate the preference ranking of outcomes by an opponent (cf. [4]). For example, negotiation strategies that aim at maximizing an opponent's utility value by means of walking on an utility iso-curve in one's own preference profile only need adequate information about an opponent's ranking of outcomes. That is, it is sufficient when using such strategies to possess accurate ordinal ranking information.

To estimate the distance between the rankings of the bids given the actual utility function of the opponent and the learned utility function, a metric is introduced that compares all outcomes in the outcome space pairwise. In order to do so, a ranking relation \prec_U is defined as follows: $\forall \omega_i, \omega_j \in \Omega, \omega_i \prec_U \omega_j \Leftrightarrow U(\omega_i) < U(\omega_j)$. Using this ranking relation, we can define a conflict indicator function adapted from [6] to measure conflicting rankings given arbitrary utility functions u and \tilde{u} . The conflict indicator function is defined as follows:

$$c_{\prec_u, \prec_{\tilde{u}}}(\omega_i, \omega_j) = \begin{cases} 1 & \text{if } (\omega_i \preceq_u \omega_j \wedge \omega_j \prec_{\tilde{u}} \omega_i) \vee (\omega_i \prec_u \omega_j \wedge \omega_j \preceq_{\tilde{u}} \omega_i) \\ & \vee (\omega_i \preceq_{\tilde{u}} \omega_j \wedge \omega_j \prec_u \omega_i) \vee (\omega_i \prec_{\tilde{u}} \omega_j \wedge \omega_j \preceq_u \omega_i), \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The conflict indicator function yields 1 when the ranking relation of two arbitrary outcomes ω, ω' based on the learned utility space \tilde{U} is not the same as the ranking relation based on the actual utility space of the opponent U ; if the rankings based on both utility functions match the conflict indicator takes the value of 0.

Using the conflict indicator c , we can define a metric called the *ranking distance* of two utility functions. The ranking distance is the calculated average of the number of conflicts between two utility functions given c :

$$d_{ranking}(U, \tilde{U}) = \frac{1}{|\Omega|^2} \sum_{\omega \in \Omega, \omega' \in \Omega} c_{\prec_U, \prec_{\tilde{U}}}(\omega, \omega') \quad (3)$$

In [6] various properties of this distance measure are proved, including e.g. reflexivity, symmetry and the triangle inequality property.

It is useful to not only apply the distance measures to complete preference profiles but also to apply it to the issue priorities or weights in such a profile. In Section 4 we apply the assessment method to the learning approach for automated closed negotiation based on Bayesian learning proposed in [8]. In this learning approach the different components of a linearly additive utility function, i.e. weights and evaluation functions, are learned in a different way. In order to obtain experimental data about these different learning processes we therefore also define similar distance measures to those discussed above for measuring distance of actual and learned issues weights.

The set of weights can be represented as a weight vector, and it is not hard to define the Pearson correlation coefficient for the vectors of weights. The coefficient is defined as follows:

$$d_{pearson}(W, \tilde{W}) = \frac{\sum_{i=1}^n (w_i - \langle w \rangle) * (\tilde{w}_i - \langle \tilde{w} \rangle)}{\sqrt{\sum_{i=1}^n (w_i - \langle w \rangle)^2 \sum_{i=1}^n (\tilde{w}_i - \langle \tilde{w} \rangle)^2}} \quad (4)$$

To calculate the ranking distance between the two weight vectors W and \tilde{W} a ranking relation is constructed on the weights of the corresponding vector as follows: $i = 1 \dots n, j = 1 \dots n, i \prec j \Leftrightarrow w(i) < w(j)$, where $w(i) = w_i$. Using this relation, the conflict indicator $c_{\prec_W, \prec_{\tilde{W}}}(i, j)$ then can be defined in the same way as for utility functions. The ranking distance of two weight vectors is defined as follows:

$$d_{ranking}(W, \tilde{W}) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n c_{\prec_W, \prec_{\tilde{W}}}(i, j) \quad (5)$$

3.2 Negotiation Domains and Profiles

Whereas precise mathematical metrics can be defined for measuring distance of preference profiles, for the selection of an adequate set of domains to be used in the experimental setup less formal criteria are proposed here. The main reason is that it is impossible to assess a learning technique on the space of all negotiation domains and associated preference profile. Ideally, then, one would use an experimental setup based on random sampling of the domains and profiles in order to deal with this problem. However, it is not clear how to setup such a sampling procedure.¹ Instead, we therefore discuss and propose to use three factors for selecting domains that are relevant in testing the learning quality.

Size of the negotiation domain. The amount of information exchanged during the negotiation is limited in a closed negotiation since we can rely only on observed

¹ As an example, we found that the predictability of issue preferences (see below) may influence the outcomes of negotiation strategies. It is not particularly clear, however, how to obtain a random sample which would be an adequate representation of domains with and without predictable issues.

negotiation moves of an opponent, which affects learning quality. The amount of information needed by a learning technique typically depends on the model structure and the size of the parameter space that is to be learned. Therefore, a learning technique has to be assessed on negotiation domains of various sizes and of various complexity. Since in any negotiation the number of issues is one of the most important factors that determines the complexity of the preferences profile, a set of domains should be selected that range from a low number of issues to higher number of issues.

Predictability of the preferences. Most learning techniques for learning an opponent's preference profile use assumptions about the structure of the preference profile (e.g. see [2, 8, 22]). Among others such techniques may rely on the predictability of issue preferences [7]. Issues are called predictable when even though the actual evaluation function for the issue is unknown, it is possible to guess some of its global properties. For example, a price issue typically is rather predictable, where more is better for the seller, and less is better for the buyer, and the normal ordering of the real numbers is maintained; an issue concerning colour, however, is typically less predictable. Learning even ranking preferences related to issue values of unpredictable issues therefore is more difficult.

The set of selected negotiation domains for any experimental setup therefore ideally should consist of a balanced mix of predictable and unpredictable issues. In principle, the higher the number of unpredictable issues the more complicated the learning of a corresponding profile becomes.

Opposition of preferences. The results of analyzing negotiation dynamics presented in [7] revealed that some negotiation strategies are sensitive to preference profiles with compatible issues. Issues are compatible if the issue preferences of both negotiating parties are such that they both prefer the same alternatives for the given issue. Negotiation strategies may more or less depend on whether preferences of the negotiating parties are opposed or not on every issue. That is, using some strategies it is harder or even impossible to exploit such common ground and agree on the most preferred option by both parties for compatible issues (humans are reported to have difficulty with this as well; cf. [19]). A selection of preference profiles should therefore take into account that both preference profiles with and without compatible issues are included.

The notion of opposition can be made more precise. In [10] a notion of local opposition based on the gradients of the utility functions of both parties is defined for each outcome in the negotiation domain. Intuitively, if the gradients point to opposite directions then the preferences of the negotiation parties are opposed. The more colinear the gradients are the closer (more compatible) the preferences of the parties. Although it is possible to generalize the notion of local opposition relative to an outcome to a more global notion of opposition of utility functions, we propose to reuse the distance measures for preference profiles to measure the level of opposition present.

As discussed, it is not clear how to randomly sample negotiation domains and we use the criteria discussed to select a number of negotiation domains to be used

in our experimental setup. The selection we present is not intended to cover all variations in line with these criteria but rather is ment to illustrate these criteria. The following negotiation domains with predefined opponent profiles have been selected (also see Table 1 for details about profile distances):

- Second hand car selling, taken from [9]: a domain of 5 issues, of which only price is really predictable. That is, an agent can only reliably predict the other agent’s preferences for this issue.
- Service-Oriented Negotiation, taken from [2], a domain with 4 issues for which domain knowledge is made available to the strategies.
- Employment contract negotiation domain, taken from [13] with 5 discrete issues. All issues have predictable values. The preference profiles have the strongest opposition in our setup.
- AMPO vs City, taken from [15], a domain with 10 issues, for which 7 are rather predictable, but 3 are not. This is the biggest domain in our experimental setup.
- Party domain is created for negotiation experiments with humans. It is a rather small domain with 5 discrete issues with 5 possible values each. All of the issues are unpredictable. The preference profiles have the lowest opposition in the experimental setup.

Domain	Utility spaces		Weights		Domain size	No. of Predictable
	Ranking	Pearson	Ranking	Pearson		
AMPO vs. City	0.662	-0.482	0.422	-0.139	7,128,000	3 (10)
Party	0.540	-0.126	0.467	-0.276	3,125	0 (5)
SON	0.669	-0.453	0.833	-0.751	810,000	4 (4)
Employment contract	0.698	-0.584	0.600	-0.241	3,125	5 (5)
2nd hand car	0.635	-0.387	0.600	-0.147	18,750	1 (5)

Table 1. Distance measures between utility space in the analyzed domains

3.3 Negotiation Strategies of the Opponent

The results of the analysis presented in [7] also have shown that the performance of a negotiation strategy can be significantly influenced by the negotiation strategy of the opponent. For example, the class of pure time-dependent tactics (TDT; see [3]) does not take into account the negotiation moves of opponents and selects the next offer to propose in a negotiation based on how close one is to the negotiation deadline. Whereas TDT tactics are insensitive to opponent moves, negotiation strategies in the class of behaviour-dependent tactics (BDT) do base their choice of offer on the offers received so far from the opponent. A variety of strategies therefore is needed to asses the quality of learning, which includes strategies that belong to the TDT class, the BDT class as well as mixes thereof.

The selection of strategies to be used in an experimental setup should be able to test the robustness of the learning technique with respect to various opponents that use different types of negotiation strategies. For example, to

enable learning of opponent preferences from the observed negotiation moves (offers) typically a concession assumption is made (cf. [1, 8, 22]). Such rationality assumptions might however be exploited and it should be tested if a learning technique is robust against strategies like the Zero-Intelligence strategy that uses an irrational random tactic [5].

Again we do not claim to present an exhaustive coverage of the criteria discussed, but present a selection to illustrate. The following negotiation strategies have been used by the negotiating parties in our experimental setup:

- The ABMP strategy from [9], which is a concession oriented approach in the TDT class, and is taking no heed of knowledge about the domain or the opponent. The ABMP strategy uses a non-linear concession tactic. It concedes more in the beginning of the negotiation when the gap between the opponents' negotiation positions is big and decreases the size of the concession when their negotiation positions approach each other. As such, it is an example of a so-called conceder tactic (cf. [3]).
- The Trade-off Strategy, taken from [4], uses so-called similarity criteria and exploits domain knowledge. The Trade-off strategy is an example of a Behaviour-dependent strategy. In our experiments we allowed three smart steps and a concession of 0.05 for the smart meta strategy.
- Zero-Intelligence, taken from [5], is a random strategy that makes random jumps through the outcome space. The ZI agent used a reservation point in our experiments to avoid making offers that have very low utility which was set to 0.6. The ZI strategy plays a role as a baseline strategy.

4 Application and Experimental Results

To show how the proposed method is used in practice we apply it to agents that make use of opponent's preferences learning techniques taken from [8]. This section presents the details of the experimental setup and presents some results obtained.

4.1 Experimental Setup

To test the proposed method for learning quality analysis an experimental setup is created. A learning technique based on Bayesian learning algorithm proposed in [8] is used as a subject of the analysis. The opponent model in [8] is based on learning probability over a set of hypothesis about evaluation functions and weights of the issues. The probability distribution is defined over the set of hypothesis that represent agent's belief about opponent's preferences. Structural assumptions about the evaluation functions and weights are made to decrease the number of parameters to be learned and simplify the learning task.

Authors propose two versions of the learning algorithm. In the first version of the algorithm each hypotheses represents a complete utility space as a combination of weights ranking and shapes of the issue evaluation functions. The

size of the hypothesis space growth exponentially w.r.t. the number of issue and thus is intractable for negotiation domains with high number of issues.

The second version of the algorithm is a scalable variant for the first one. In this version the agent tries to learn probability distribution over the individual hypothesis about the value of the weight and shape of the issue evaluation function independently of other issues. The computational tractability of the learning is achieved by approximating the conditional distributions of the hypotheses using the expected values of the dependant hypotheses.

The authors realize that the learning power of such solution would degrade compared to the first version of the algorithm. However, the agent performs quite well on the negotiation domains of higher dimensionality (10 issues), see [8]. It is only reasonable to expect that the quality of learning degrades when the size of the negotiation domain is increased. In addition, the unscaleable version of the Bayesian learning agent is expected to perform less than the scalable version in terms of learning quality on the same negotiation domain.

The ZI agent is used to test the robustness of the Bayesian learning technique with respect to the opponent's negotiation strategy. Both versions of the Bayesian learning algorithm are expected to show worse performance when negotiating against the ZI agent because they rely on assumptions about the rationality of the opponent's negotiation strategy. Furthermore, the quality of learning of the Bayesian agent on the Party domain is expected to suffer from the unpredictability of the preference of all the issues in this domain.

4.2 Evaluation

Due to space limitations, we only present the results of those experiments that give rise to some of the more significant conclusions. Figure 1 shows results that represent the quality of learning of the scalable version of the Bayesian agent on the Employment contract negotiation domain.

The results show that the concession tactics of the opponent influences the quality of learning, as expected. The Bayesian agent learns the preferences of the opponent better when the opponent uses the Trade-off tactics rather than the ABMP strategy. The Trade-off tactics uses semi-linear concession tactics (see previous section), which is more consistent with the opponent tactics assumptions made in the Bayesian agent. As expected, the Bayesian agent learns the opponent preference slower in case of the ZI negotiation strategy of the opponent. However, it is still capable of learning the opponent's preference quite well.

In general, the learning quality is better in smaller negotiation domains spaces. This follows from a comparison of the AMPO vs. City domain which is the largest domain with the SON contract negotiation domain in Figure 2. The Bayesian learning technique is able to perfectly learn a model of the ranking (ranking distance measure) in case of the SON domain and keeps improving the absolute values of the weights (cf. the results of the Pearson distance). In the AMPO vs. City domain the agent is able to learn the ranking of the weights to some extent. However, the results show that the learning of the outcome ranking

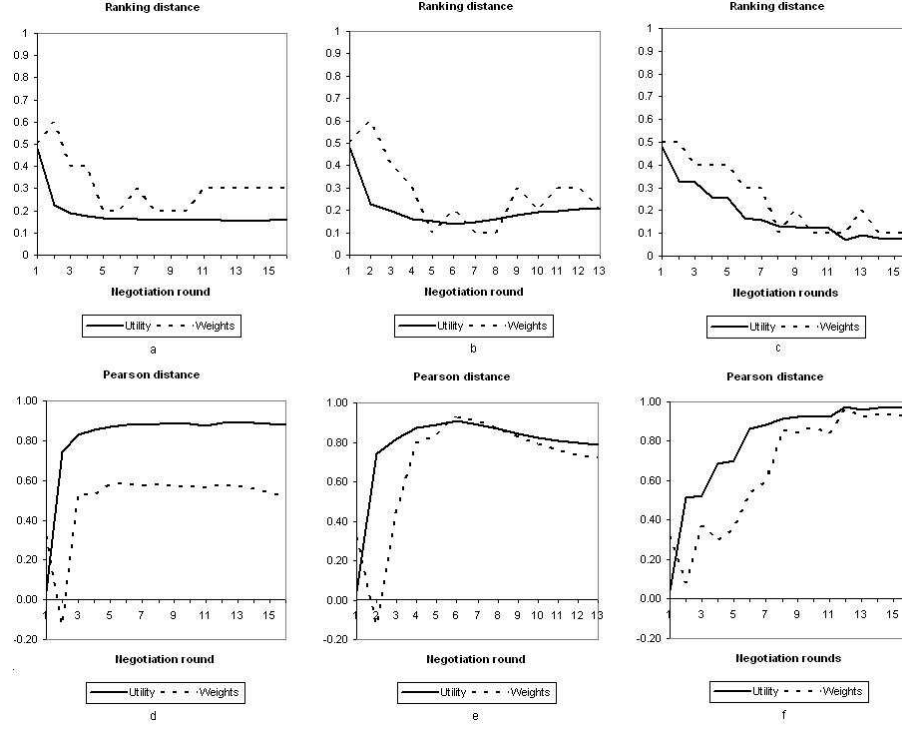


Fig. 1. Learning quality measures for the unscaleable Bayesian agent on the Employment contract negotiation domain against the Trade-off (a,d), ABMP (b,e), and ZI (c,f) strategies.

is rather limited. This can be partly explained by the presense of a few issues with unpredictable preferences, which results in a lower learning quality of the evaluation functions and a lower learning quality of the utility function compared to the results for the SON domain.

Another interesting observation is that the learning algorithm approximates the absolute, cardinal values of the utility function and weights quite well, see Figure 2(b,e). This can be explained by the nature of the hypothesis space of the learning algorithm. The algorithm calculates opponent's utility values of a bid as an expected value of a random variable. The expected value is a sum of the utilities according to the hypothesis weighted according to their probabilities. Thus, even if the more detailed structure of the opponent's preferences is not learned by the agent the information learned can still be used to approximate the utility function of the opponent as a linear combination of the set of all hypotheses.

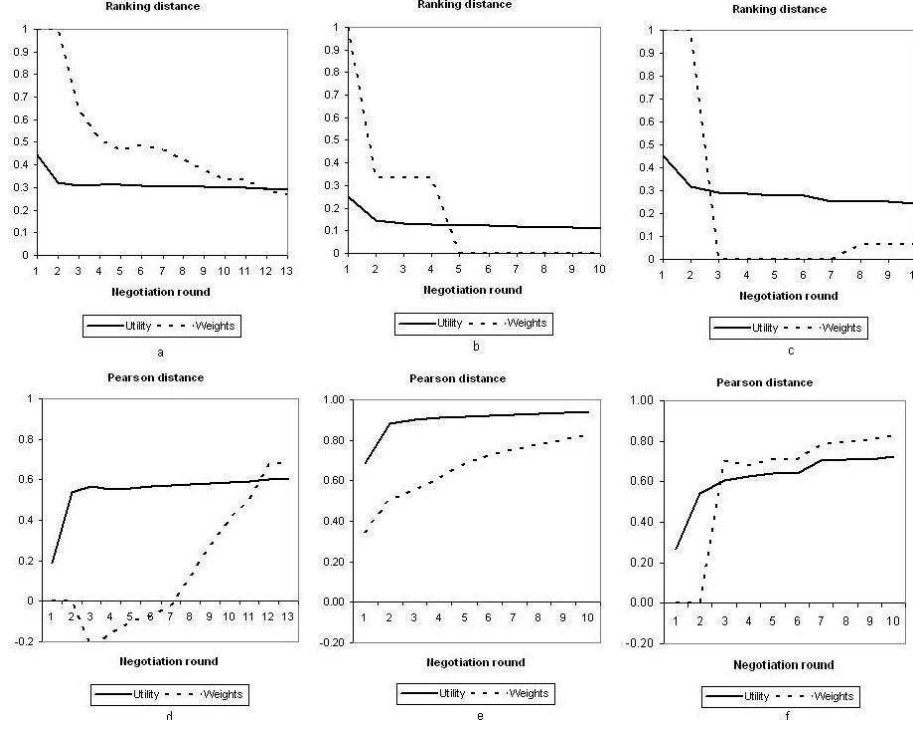


Fig. 2. Learning quality measures for the scaleable Bayesian agent against Trade-off strategy on the AMPO vs. City (a,d), SON (b,e), and Party (c,f) negotiation domains.

5 Conclusion and Discussion

In this paper, a method for the analysis of the learning quality of learned opponent preference profiles in automated negotiation has been presented. The method consists of three components: (i) It uses distance measures between the actual preference profile of the opponent and the learned preference profile to assess the quality of the learned model; (ii) it propose criteria for the systematic classification of negotiation domains and preferences profiles to assess the impact of a variety of domains on the quality of the learned model; and (iii) it proposes some criteria to select a set of negotiation strategies. The method has been applied as an illustration to the agents introduced in [8], which use a Bayesian learning technique to learn preference profiles. The results confirmed the effectiveness of the learning technique, in particular that of the unscaleable version of the agent. The performance of the scalable agent is comparable to that of the unscaleable version on smaller negotiation domains. As is to be expected, the results showed that learning performance slowly degrades for the negotiation domains of higher dimensionality.

In addition, we believe that the results revealed interesting features of the learning algorithms that can be used to improve their performance. For instance, a more detailed analysis of the unscaleable version of the Bayesian agent proposed in [8] revealed that the learning algorithm typically approximates the opponent's preference profile with a weighted sum of the evaluation hypotheses instead of learning the evaluation function that best matches the actual function defined for a given issue. This indicates that it may be possible to reduce the hypothesis space of evaluation functions and use a smaller set of such functions which would still be sufficient to approximate a wide range of possible opponent preference profiles.

One issue that needs more research is the classification of both domains as well as strategies in order to systematically evaluate a learning technique used in negotiation. Though in principle, given that utility functions representing preference profiles are of a certain type, it is possible to generate random outcome and utility spaces to be used for evaluation. A method based on such random samples would potentially be more generic than the analysis of cases as used in this paper. It is not trivial to define such an approach, however, since the distribution used to generate samples should not favour particular situations (e.g. combinations of strictly opposed preference profiles over compatible preference profiles). Since it is also not completely clear how realistic certain preference profiles which lack sufficient structure are this poses another problem to be addressed before the use of random samples would make sense.

In this paper we have mainly focussed on the learning of the preference profile of the opponent. However, a negotiation strategy might also exploit other knowledge about the negotiating parties. Other parameters such as reservation value, or type of negotiation strategy however would require certain adaptations of the proposed method. We plan to extend our method with such tools along with a statistical toolbox useful in defining a testbed for automated negotiation agents.

References

1. S. Buffett and B. Spencer. Learning opponents preferences in multi-object automated negotiation. In *Seventh International Conference on Electronic Commerce (ICEC'05)*, pages 300–305. New York: ACM, 2005.
2. R. Coehoorn and N. Jennings. Learning an opponents preferences to make effective multi-issue negotiation trade-offs. In *Proceedings of 6th International Conference on E-Commerce*, pages 59–68, 2004.
3. P. Faratin, C. Sierra, and N. R. Jennings. Negotiation decision functions for autonomous agents. *Int. Journal of Robotics and Autonomous Systems*, 24(3-4):159–182, 1998.
4. P. Faratin, C. Sierra, and N. R. Jennings. Using similarity criteria to make negotiation trade-offs. *Journal of Artificial Intelligence*, 142(2):205–237, 2003.
5. D. K. Gode and S. Sunder. Allocative efficiency in markets with zero intelligence (zi) traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.

6. V. Ha and P. Haddawy. Similarity of personal preferences: Theoretical foundations and empirical analysis. *Artificial Intelligence*, 146(2):149–173, 2003.
7. K. Hindriks, C. Jonker, and D. Tykhonov. Negotiation dynamics: Analysis, concession tactics, and outcomes. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'07)*, pages 427–433, 2007.
8. K. Hindriks and D. Tykhonov. Opponent modelling in automated multi-issue negotiation using bayesian learning. In *Proceedings of the AAMAS 2008*, 2008.
9. C. M. Jonker, V. Robu, and J. Treur. An agent architecture for multi-attribute negotiation using incomplete preference information. *Journal of Autonomous Agents and Multi-Agent Systems*, 15(2):221–252, 2007.
10. G. E. Kersten and S. J. Noronha. Rational agents, contract curves, and inefficient compromises report. Technical report, International Institute for Applied Systems Analysis, 1997.
11. R. Lin, S. Kraus, J. Wilkenfeld, and J. Barry. An automated agent for bilateral negotiation with bounded rational agents with incomplete information. In *Proceedings of the International European Conference on AI (ECAI'06)*, pages 270–274, 2006.
12. W. W. H. Mok and R. Sundarraj. Learning algorithms for single-instance electronic negotiations using the time-dependent behavioral tactic. *ACM Transactions on Internet Technology*, 5(1):195–230, February 2005.
13. J. Nadler, L. Thompson, and L. van Boven. Learning negotiation skills: Four models of knowledge creation and transfer. *Journal of Management Science*, 49(4):529–540, 2003.
14. V. Narayanan and N. Jennings. Learning to negotiate optimally in non-stationary environments. In *Proceedings of the International Conference on Cooperative Information Agents (CIA'06)*, volume 4149 of *LNAI*, pages 288–300, 2006.
15. H. Raiffa. *The Art and Science of Negotiation*. Harvard University Press, 1982.
16. H. Raiffa, J. Richardson, and D. Metcalfe. *Negotiation Analysis: The Science and Art of Collaborative Decision Making*. Harvard University Press, 2003.
17. A. Restificar and P. Haddawy. Inferring implicit preferences from negotiation actions. In *Proceedings of the International Symposium on Artificial Intelligence and Mathematics*, 2004.
18. A. G. M. Robert H. Guttman and P. Maes. Agent-mediated electronic commerce: a survey. *The Knowledge Engineering Review*, 1998.
19. L. Thompson. *The Mind and Heart of the Negotiator*. Prentice Hall, 2004.
20. J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.
21. D. Zeng and K. Sycara. Benefits of learning in negotiation. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*, 1997.
22. D. Zeng and K. Sycara. Bayesian learning in negotiation. *International Journal of Human Computer Systems*, 48:125–141, 1998.