Engineering Applications of Artificial Intelligence 34 (2014) 45–57

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

An efficient and robust negotiating strategy in bilateral negotiations over multiple items

Jianye Hao [a,d,*], Songzheng Song [b], Ho-fung Leung [c], Zhong Ming [d,**]

[a] *Massachusetts Institute of Technology, United States*
[b] *Nanyang Technological University, Singapore*
[c] *The Chinese University of Hong Kong, Hong Kong Special Administrative Region*
[d] *Shenzhen University, China*

ARTICLE INFO

Article history:
Received 29 December 2013
Received in revised form
21 April 2014
Accepted 8 May 2014
Available online 3 June 2014

Keywords:
Bilateral negotiation
Multiagent learning
Robustness
Model checking

ABSTRACT

Multi-item negotiations surround our daily life and usually involve two parties that share common or conflicting interests. Effective automated negotiation techniques should enable the agents to adaptively adjust their behaviors depending on the characteristics of their negotiating partners and negotiation scenarios. This is complicated by the fact that the negotiation agents are usually unwilling to reveal their information (strategies and preferences) to avoid being exploited during negotiation. In this paper, we propose an adaptive negotiation strategy, called *ABiNeS*, which can make effective negotiations against different types of negotiating partners. The *ABiNeS* strategy employs the non-exploitation point to adaptively adjust the appropriate time to stop exploiting the negotiating partner and also predicts the optimal offer for the negotiating partner based on the reinforcement-learning based approach. Simulation results show that the *ABiNeS* strategy can perform more efficient exploitations against different types of negotiating partners, and thus achieve higher overall payoffs compared with the state-of-the-art strategies under negotiation tournaments. We also provide a detailed analysis of why the *ABiNeS* strategy can negotiate more efficiently compared with other existing state-of-the-art negotiation strategies focusing on two major components. Lastly, we propose adopting the *single-agent best deviation* principle to analyze the robustness of different negotiation strategies based on model checking techniques. Through our analysis, the *ABiNeS* strategy is shown to be very robust against other state-of-the-art strategies under different negotiation contexts.

1. Introduction

Negotiation is a common and important approach to resolve conflicts and reach agreements between different parties in our daily life. With the advance and popularity of Web and e-commerce, it is expected that a lot of previous negotiation activities between humans will be moved to electric platforms and greatly benefit from automated negotiation techniques (Kersten and Noronha, 1999). Automated negotiation techniques can, to a large extent, alleviate the efforts of human negotiators, and also aid human in reaching better negotiation outcomes by compensating for the limited computational abilities of humans when they are faced with complex negotiations. Until now, a lot of automated negotiation strategies and mechanisms have been proposed in different negotiation scenarios (Faratin et al., 2003; Saha et al., 2005; Hindriks and Tykhonov, 2008; Jakub and Ryszard, 2006).

The major difficulty in designing automated negotiation agent is how to achieve optimal negotiation results given incomplete information on the negotiating partner. The negotiation partner usually keeps its negotiation strategy and its preference as its private information to avoid exploitations. A lot of research effort has been devoted to better understand the negotiation partner by either estimating the negotiation partner's preference profile (Zeng and Sycara, 1998; Hindriks and Tykhonov, 2008; Coehoorn and Jennings, 2004) or predicting its decision function (Zeng and Sycara, 1996; Jakub and Ryszard, 2006). On one hand, with the aid of different preference profile modeling techniques, the negotiating agents can get a better understanding of their negotiating partners and thus increase their chances of reaching mutually beneficial negotiation outcomes. On the other hand, effective strategy prediction techniques enable the negotiating agents to maximally exploit their negotiating partners and thus receive as much benefit as possible from negotiation. However, in most of previous work, the negotiating agents are usually assumed to be situated in a

* Corresponding author at: Massachusetts Institute of Technology, United States.
** Corresponding author.
E-mail addresses: jianye@mit.edu (J. Hao), songsz@nut.edu.sg (S. Song), lhf@cuhk.edu.hk (H.-f. Leung), mingz@szu.edu.cn (Z. Ming).

http://dx.doi.org/10.1016/j.engappai.2014.05.008
0952-1976/

negotiation environment that strong limitations are put on the negotiation scenario or the negotiation opponent. For example, some work assumes (Saha et al., 2005) that the agents negotiate over a single item only, however practical negotiation scenarios usually involve multiple items to negotiate over. There also exists some work that assumes that the negotiation agents can have access to their opponents' (partial) preferences. This can be unrealistic especially in multi-issue negotiation scenarios in which the preferences of different agents may vary significantly, and the agents usually would not like to disclose their preferences to avoid being exploited. Another assumption commonly adopted (Jakub and Ryszard, 2006) is that the negotiation opponent is limited to choose from a specific set of simple strategies, e.g., time-dependent or behavior-dependent tactics. Those strategies designed under this assumption may not work well against other negotiation partners with more complex state-of-the-art strategies.

To this end, in recent years a number of advanced negotiation strategies taking advantage of existing techniques have been proposed and agents employing these strategies have participated in *automated negotiating agents competition* (*ANAC*) (Baarslag et al., 2010, 2013). The ANAC competition provides a negotiation platform which enables different negotiation agents to be evaluated against a wide range of opponents within a realistic negotiation environment. During the past three years, dozens of state-of-the-art negotiation strategies have been extensively evaluated in a variety of multi-issue negotiation scenarios and valuable insights have been obtained in terms of the advantages and disadvantages of different techniques, e.g., the efficacy of different acceptance conditions (Baarslag et al., 2011). It is still an open and interesting problem to design more efficient automated negotiation strategies against a variety of negotiating opponents in different negotiation domains.

In this paper, we propose an adaptive negotiation strategy *ABiNeS* for automated agents to negotiate in bilateral multi-issue negotiation environments following the settings adopted in ANAC 2012 (The Third International Automated Negotiating Agent Competition, 2012). Bilateral multi-issue negotiations surround people's daily life and have received lots of attention in the negotiation literature. During negotiation, both the agents' negotiation strategies and preference profiles are their private information, and for each agent the only available information about the negotiating partner are its past negotiation moves. Considering the diversity of the available negotiation strategies that the negotiating agents can adopt, it is usually very difficult (or impossible) to predict which specific strategy the negotiating partner is using based on this limited information. To effectively cope with different types of opponents, we introduce the concept of non-exploitation point λ to adaptively adjust the degree that an *ABiNeS* agent exploits its negotiating opponent. The value of λ is determined by the characteristics of the negotiation scenario and the concessive degree of the negotiating partner, which is estimated based on the negotiation history. Besides, to maximize the possibility that the offer the *ABiNeS* agent proposes will be accepted by its negotiating partner, it can be useful to make predictions on the preference profile of the negotiating partner. Instead of explicitly modeling the negotiation partner's preference profile, we propose a reinforcement-learning based approach to determine the optimal proposal for the negotiating partner based on the current negotiation history.

We evaluate the performance of the *ABiNeS* strategy compared with a number of state-of-the-art negotiation strategies from two different perspectives: *efficiency* in terms of the average payoff obtained under a particular negotiation tournament and *robustness* in terms of how likely the agents have the incentive to adopt our strategy rather than other strategies. First, the *efficiency* evaluation is conducted under the negotiation tournament setting following ANAC 2012 using GENIUS[1] (Lin et al., 2012) platform. Simulation results show that the *ABiNes* strategy can make more effective exploitations against a variety of negotiation partners and thus obtain higher average payoffs during negotiation tournaments and it is worth mentioning that the *ABiNes* strategy wins the champion of ANAC 2012 known as *CUHKAgent*. Second, we give a detailed analysis of the *ABiNes* strategy by investigating the influence of its two major novel components. Through the detailed analysis, we aim at providing a clear understanding of why the *ABiNes* strategy can win the champion of ANAC 2012, and more importantly, offering valuable insights for the automated negotiation community for the future negotiation strategy design. Third, we propose adopting the *single-agent best deviation* principle to analyze the robustness of different negotiation strategies based on model checking techniques. According to the *robustness* analysis, the *ABiNes* strategy is shown to be very robust against other state-of-the-art strategies under different negotiation contexts.

The remainder of the paper is organized as follows. In Section 2, we give a description of negotiation model we consider in this paper. In Section 3, the negotiation strategy *ABiNes* we propose is introduced. In Section 4, we give detailed evaluation of the negotiation *efficiency* and *robustness* of *ABiNes* compared with the state-of-the-art negotiation strategies under different negotiation contexts. An overview of related work on automated negotiation strategies is given in Section 5. Lastly conclusion and future work are given in Section 6.

2. Negotiation model

In this section, we describe the negotiation model we consider in this work, which follows the settings adopted in ANAC 2012 (The Third International Automated Negotiating Agent Competition, 2012). We focus on bilateral negotiations, i.e., negotiations between two agents. Specifically, the alternating-offers protocol is adopted to regulate the interactions between the negotiating agents, in which the agents take turns to exchange proposals. For each negotiation scenario, both agents can negotiate over multiple issues (items), and each item can have a number of different values. Let us denote the set of items as $\mathcal{M}$, and the set of values for each item $m_i \in \mathcal{M}$ as $\mathcal{V}_i$.[2] We define a negotiation outcome ω as a mapping from every item $m_i \in \mathcal{M}$ to a value $v \in \mathcal{V}_i$, and the negotiation domain is defined as the set Ω of all possible negotiation outcomes. For each negotiation outcome ω, we use $\omega(m_i)$ to denote the corresponding value of the item m_i in the negotiation outcome ω. We assume that the knowledge of the negotiation domain is known to both agents beforehand, and is not changed during the whole negotiation session.

For each negotiation outcome ω, different agents may have different preferences. Here we assume that each agent i's preference can be modeled by a utility function u_i such that $\forall \omega \in \Omega$, it is mapped into a real-valued number in the range of [0,1], i.e., $u_i(\omega) \in [0, 1]$. In practical negotiation environments, it is usually associated with certain cost in each negotiation. To take this factor into consideration, a real-time deadline is imposed on the negotiation process and each agent's actual utilities over the negotiation outcomes are decreased by a discounting factor δ over time. Following the setting adopted in ANAC'12, each negotiation session is allocated 3 min, which is normalized into the range of [0,1], i.e., $0 \le t \le 1$. Formally, if an agreement is reached at time t before the deadline, each agent i's actual utility function $U_i^t(\omega)$ over this

[1] GENIUS is short for General Environment for Negotiation with Intelligent multi-purpose Usage Simulation.

[2] Here $\mathcal{V}_i$ can be either discrete values or continuous real values.

mutually agreed negotiation outcome ω is defined as follows:

$$U_i^t(\omega) = u_i(\omega)\delta^t \tag{1}$$

If no agreement is reached by the deadline, each agent i will obtain a utility of $ru_i^0\delta$, where ru_i^0 is agent i's private reservation value in the negotiation scenario. The agents will also obtain their corresponding reservation values if the negotiation is terminated before the deadline. Note that the agents' actual utilities over their reservation values are also discounted by the discounting factor δ over time t. We assume that the agents' preference information and their reservation values are private and cannot be accessed by their negotiating partners.

The interaction between the negotiation agents is regulated by the alternating-offers protocol, in which the agents are allowed to take turns to exchange proposals. During each encounter, if it is agent i's turn to make a proposal, it is allowed to make a choice from the following three options:

Accept the offer from its negotiating partner: In this case, the negotiation ends and an agreement is reached. Both agents will obtain the corresponding utilities according to Eq. (1), where ω is the negotiation outcome that they mutually agree with.

Reject and propose a counter-offer to its negotiating partner: In this case, the negotiation process continues and it is its negotiating partner's turn to make a counter-proposal next time provided that the deadline is not reached yet.

Terminate the negotiation: In this case, the negotiation terminates and each agent i gets its corresponding utility based on its private reservation value with the initial value of ru_i^0. Note that their actual utilities in this case are also decreased over time by the same discounting factor δ, i.e., $U_i^t = ru_i^0 * \delta^t$. Note that the reservation value may be different for different negotiation parties and also vary in different negotiation domains.

Overall, the negotiation process terminates when either of the following conditions is satisfied: (1) the deadline is reached (*End*); (2) an agent chooses to terminate the negotiation before reaching deadline (*Terminate*); (3) an agent chooses to accept the negotiation outcome proposed by its negotiating partner (*Accept*).

For each negotiation session between two agents A and B, let $x_{A\to B}^t$ denote the negotiation outcome proposed by agent A to agent B at time t. Naturally a negotiation history $H_{A\leftrightarrow B}^t$ between agent A and B until time t can be represented as follows:

$$H_{A\leftrightarrow B}^t := (x_{p_1\to p_2}^{t_1}, x_{p_3\to p_4}^{t_2}, \ldots, x_{p_n\to p_{n+1}}^{t_n}) \tag{2}$$

where

- $t_k \le t_l$ for $k \le l$, i.e., the negotiation outcomes are ordered over time, and also $t_n \le t$.
- $p_k = p_{k+2} \in \{A, B\}$, i.e., the negotiation process strictly follows the alternating-offers protocol.

Similarly, we denote the negotiation history during a certain period of time (between time t_1 and t_2) as $H_{A\leftrightarrow B}^{t_1\to t_2}$. As previously mentioned, we know that a negotiation session between two agents A and B will terminate either when one agent chooses an action from the set $\{Accept, Terminate\}$, or the deadline is reached. Thus the last element of a complete negotiation history can be either one of the elements from the set $S = \{Accept, Terminate, End\}$. A negotiation history by time t is active if its last element is not equal to any element in S.

3. *ABiNeS*: an adaptive bilateral negotiating strategy

In this section, we describe the adaptive negotiating strategy *ABiNeS* in detail. For the ease of description, we refer to the *ABiNeS* agent as agent A and its negotiating partner as agent B in the following descriptions. The *ABiNeS* strategy mainly consists of four basic decision components. The first component is *Acceptance-Threshold (AT)* component and it is responsible for determining the *ABiNeS* agent's minimum acceptance threshold l_A^t at time t. The second component is *Next-Bid (NB)* component whose function is to determine the negotiation outcome $x_{A\to B}^t$ that the *ABiNeS* agent proposes at time t. The third component, *Acceptance-Condition (AC)* component, is used for determining whether to accept the current proposal from agent B or not. Given a negotiation history $H_{A\leftrightarrow B}^t$, its acceptance threshold l_A^t, and its negotiation outcome $x_{A\to B}^t$ to propose at time t, the AC component returns a boolean value indicating whether to accept the offer or not, which is denoted as $AC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t)$. The last component is *Termination-Condition (TC)* component. It is responsible for determining whether to terminate the current negotiation or not. Similar to the AC component, given a negotiation history $H_{A\leftrightarrow B}^t$, its acceptance threshold l_A^t, its negotiation outcome $x_{A\to B}^t$ to propose at time t, and its reservation utility ru_A^0, the TC component returns a boolean value indicating whether to terminate the negotiation or not, which is denoted as $TC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t, ru_A^0)$.

Algorithm 1. Overall flow of the *ABiNeS* strategy.

for each negotiation history $H_{A\leftrightarrow B}^t := (x_{p_1\to p_2}^{t_1}, x_{p_3\to p_4}^{t_2}, \ldots, x_{p_n\to p_{n+1}}^{t_m})$ at current time t **do**
  Determine the acceptance threshold l^t and the negotiation outcome $x_{A\to B}^t$ to be proposed to the negotiating partner B, and $ru_A^t = ru_i^0 * \delta^t$.
  if $H_{A\leftrightarrow B}^t$ is empty **then**
    if $TC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t, ru_A^0)$ is false **then**
      Propose the negotiation outcome $x_{A\to B}^t$ to the negotiating partner.
    else
      Terminate the negotiation.
    end if
  else
    if $AC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t)$ is true and $TC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t, ru_A^0)$ is false **then**
      Accept the offer.
    else if $AC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t)$ is false and $TC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t, ru_A^0)$ is true **then**
      Terminate the negotiation.
    else if $AC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t)$ is true and $TC(H_{A\leftrightarrow B}^t, l_A^t, x_{A\to B}^t, ru_A^0)$ is true **then**
      if $U_A^t(x_{p_n\to p_{n+1}}^{t_m}) > ru_A^t$ **then**
        Accept the offer.
      else
        Terminate the negotiation
      end if
    else

21: Propose the negotiation outcome $x^{t}_{A\rightarrow B}$ to the negotiating partner.
22: **end if**
23: **end if**
24: **end for**

The overall description of the negotiation strategy *ABiNeS* based on the above basic elements is presented in Algorithm 1. At time t the acceptance threshold and the next-round negotiation outcome are calculated first based on the AT and NB components respectively. If the *ABiNeS* agent is the first one to make a proposal, then it is faced with two options: either proposing a negotiation outcome to agent B or terminating the negotiation (Lines 2–8). Otherwise, the *ABiNeS* agent makes a choice among three options: proposing a negotiation outcome to agent B, choosing to terminate the negotiation and accepting the offer from agent B (Lines 10–22). The decisions on whether to accept the offer or terminate the negotiation are determined by the AC and TC components respectively. We introduce each decision component in detail in the following sections.

3.1. Acceptance-threshold (AT) component

The AT component is responsible for determining the acceptance threshold l of the *ABiNeS* strategy during negotiation. The value of the acceptance threshold reflects the agent's current concession degree and should be adaptively adjusted based on the past experience and the characteristic of the negotiation environment. Besides, it also has explicit influence on the decision-making process of the other three components, which will be introduced later.

We assume that the negotiating partner is myopic self-interested, and it will accept any proposal when the deadline is approaching ($t \sim 1$).[3] Therefore the acceptance threshold of an *ABiNeS* agent should be always higher than the highest utility it can obtain when $t=1$. Specifically, at any time t, the acceptance threshold l_t of the *ABiNeS* agent should not be lower than $u^{max}\delta^{1-t}$, where u^{max} is its maximum utility over the negotiation domain without discounting. Since the negotiating goal is to reach an agreement which maximize the agent's own utility as much as possible, its negotiating partner should be exploited as much as possible by setting its acceptance threshold as high as possible. On the other hand, due to the discounting effect, the actual utility the agent receives can become extremely low though its original utility over the mutually agreed negotiation outcome is high, if it takes too long for the agents to reach the agreement. In the worst case the negotiation may end up with a break-off and each agent obtains zero utility. Thus we also need to make certain compromises to the negotiating partner, i.e., lower the acceptance threshold, depending on the type of the partner we are negotiating with. Therefore, the key problem is how to balance the trade-off between exploiting and making compromise to the negotiating partner. Towards this end, we introduce the adaptive non-exploitation point λ, which represents the specific time when we should stop exploitations on the negotiating partner. This value is adaptively adjusted based on the behavior of the negotiating partner. Specifically we propose that for any time $t<\lambda$, the *ABiNeS* agent always exploits its negotiating partner (agent B) by setting its acceptance threshold to a value higher than $u^{max}\delta^{1-\lambda}$ and approaching this value until time λ according to certain pattern of behavior. After time λ, its acceptance threshold is set to be equal to $u^{max}\delta^{1-t}$ forever, and any proposal over which its utility is higher than this value will be accepted. Formally, the acceptance threshold l^t_A of an *ABiNeS* agent at time t is determined as follows:

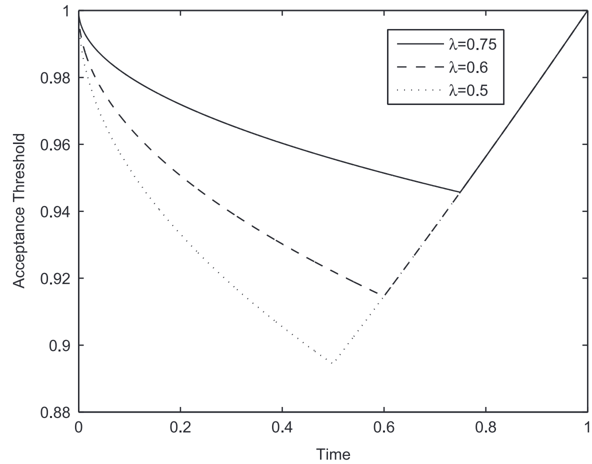

Fig. 1. The dynamics of the acceptance threshold ($u^{max}=1$, α=0.5 and δ=0.8).

$$l^t_A = \begin{cases} u^{max} - (u^{max} - u^{max}\delta^{1-\lambda})\left(\dfrac{t}{\lambda}\right)^{\alpha} & \text{if } t<\lambda \\ u^{max}\delta^{1-t} & \text{otherwise} \end{cases} \tag{3}$$

where the variable α controls the way the acceptance threshold approaches $u^{max}\delta^{1-t}$ (boulware ($\alpha>1$), conceder ($\alpha<1$) or linear ($\alpha=1$)). One example showing the dynamics of the acceptance threshold with time t with different value of λ is given in Fig. 1.

The remaining question is how to calculate the value of non-exploitation point λ. The value of λ is determined by the characteristics of the negotiation scenario (i.e., discounting factor δ) and the concession degree of the negotiating partner. The smaller the discounting factor δ is, the less actual utility we will receive as time goes by, which means more risk we are facing when we continue exploiting the negotiating partner. Therefore the value of λ should be decreased with the decreasing of the discounting factor δ. The concession degree of the negotiating partner is estimated based on its past behaviors. Intuitively, the more number of new negotiation outcomes that the negotiating partner has recently proposed, the more it is willing to make concession to end the negotiation. Specifically, the negotiation partner's concessive degree σ^t is defined as the ratio of new negotiation outcomes it proposed within the most recent finite negotiation history $H^{t'\rightarrow t}_{A\leftrightarrow B}$. If we predict that the negotiating partner is becoming more concessive, we can take advantage of this prediction by postponing the time we stop exploitations, i.e., increasing the value of λ.

Initially the value of λ is determined by the discounting factor δ only since we do not have any information on the negotiating partner yet. After that, it is adaptively adjusted based on the estimation of the concession degree of the negotiating partner. The overall adjustment rule of λ during negotiation is shown in Fig. 2.

3.2. Next-bid (NB) component

The next-bid component is responsible for determining the negotiation outcome to be proposed to the negotiating partner. Given the current acceptance threshold l^t_A at time t, any negotiation outcome over which the *ABiNeS* agent's utility is higher than l^t_A can be a reasonable outcome to propose. To maximize the

[3] Note that this is only our assumption of the opponents in the tournament negotiation setting, but in practice a negotiator might behave in the individually rational way in terms of maximizing its long-term benefit. Thus it might reject a proposal which is extremely unfair to himself to increase its relative ranking in the long run at the cost of obtaining zero utility in the current round.

Initial values

- λ_0 - the minimum value of λ,
- β - the controlling variable determining the way the value of λ changes with respect to the discounting factor δ, i.e., boulware ($\beta < 1$), conceder ($\beta > 1$) or linear ($\beta = 1$),
- σ^t - the estimation of the negotiating partner's concessive degree at time t,
- γ - the controlling variable determining the way the value of λ changes with respect to σ^t, i.e., boulware ($\gamma < 1$), conceder ($\gamma > 1$) or linear ($\gamma = 1$),
- w - the weighting factor adjusting the relative effect of σ^t on the non-exploitation point λ.

if $t = 0$ **then**

$\quad \lambda = \lambda_0 + (1 - \lambda_0)\delta^{\beta}$

end if

if $0 < t \leq 1$ **then**

$\quad \lambda = \lambda + w(1 - \lambda){\sigma^t}^{\gamma}$

end if

Fig. 2. Adjustment rule of λ at time t.

likelihood that the offer will be accepted by agent B, we need to predict the negotiation outcome ω_{max} which can maximize its utility among the set $\mathcal{C}$ of candidate negotiation outcomes (i.e., $\mathcal{C} = \{\omega | u_A(\omega) \geq l_A^t\}$). The negotiation outcome ω_{max} will be returned by the NB component as the offer to be proposed to agent B.

To obtain ω_{max}, we need to estimate agent B's private preference based on its past negotiation moves. Different approaches (Hindriks and Tykhonov, 2008; Faratin et al., 2003; Coehoorn and Jennings, 2004; Zeng and Sycara, 1998) have been proposed to explicitly estimate the negotiating partner's utility function in bilateral negotiation scenarios. To make the estimation feasible with the limited information available, we usually need to put some restrictions on the possible structures that the negotiation partner's utility function can have (Hindriks and Tykhonov, 2008) or assume that the preference profile of the negotiation partner is chosen from a fixed set of profiles (Zeng and Sycara, 1998). Instead of estimating agent B's utility function directly, here we adopt a model-free reinforcement learning based approach to predict the current best negotiation outcome for agent B. The only assumption we need here is that the negotiating partner (agent B) is individually rational and follows some kind of concession-based strategy when proposing bids, which is the most commonly used assumption in both game-theoretic approaches and negotiations (Osborne and Rubinstein, 1994; Hindriks and Tykhonov, 2008).

Based on the above assumption, it is natural to assume that the sequence of past negotiation outcomes proposed by agent B should be in accordance with the decreasing order of its preference over those outcomes. Intuitively, for a value v_i of an item m_i, the earlier and the more frequent it appears in the negotiation outcomes of the past history, the more likely that it weighs more in contributing to the negotiation partner's overall utility. Therefore, for each value of each item m_i in the negotiation domain, we keep record of the number of times that it appears in the negotiating partner's past negotiation outcomes and update its value each time a new negotiation outcome ω' is proposed by agent B as follows:

$$n(\omega'(m_i)) = n(\omega'(m_i)) + \eta^k, \quad \forall m_i \in \mathcal{M} \tag{4}$$

where ω' is the most recent negotiation outcome proposed by agent B, η is the discounting factor reflecting the decreasing speed of the relative importance of the negotiation outcomes as time increases, and k is the number of times that the value $\omega'(m_i)$ of item m_i has appeared in the history.

For each negotiation outcome ω, we define its accumulated frequency $f(\omega)$ as the criterion for evaluating the relative preference of agent B over it. The value of $f(\omega)$ is determined by the value of $n(\omega(m_i))$ for each item $m_i \in \mathcal{M}$ based on the current negotiation history. Formally, for any negotiation outcome ω, its accumulated frequency $f(\omega)$ is calculated as follows:

$$f(\omega) = \sum_{m_i} n(\omega(m_i)), \quad \forall m_i \in \mathcal{M} \tag{5}$$

The negotiation outcome ω_{max} is selected based on the ϵ-greedy exploration mechanism. With probability $1-\epsilon$, it chooses the negotiation outcome with the highest f-value from the set $\mathcal{C}$ of candidate negotiation outcomes, and chooses one negotiation outcome randomly from $\mathcal{C}$ with probability ϵ. The value of ϵ controls the exploration degree during prediction. One exception is that the negotiation outcome ω_{max} will be selected as the best negotiation outcome proposed by the negotiating agent B in the history if the NB component has received the corresponding signal from the AC component, and this will be explained in detail in the next section.

3.3. Acceptance-condition (AC) component

Given the current negotiation history $H^t_{A \leftrightarrow B}$, agent A's acceptance threshold l_A^t, and its negotiation outcome $x^t_{A \rightarrow B}$ to propose at time t, the AC component determines whether to accept the current proposal of agent B or not. The overall acceptance conditions are described in Algorithm 2. The *ABiNeS* agent accepts the proposal ω_1 from its negotiating agent B if its utility over ω_1 is either higher than its current acceptance threshold l_A^t (Lines 2 and 3). Otherwise, it checks whether there exists some negotiation outcome ω_{best} previously proposed by its negotiating agent B satisfying the above condition. If the answer is yes, then it will notify the NB component to propose ω_{best} next time (Lines 4–5).

Algorithm 2. Acceptance conditions $AC(H^t_{A\leftrightarrow B}, l^t_A, x^t_{A\rightarrow B})$.

1: Initialization: $\omega_{best} = \text{BestNegotiationOutcome}\ (H^t_{A\leftrightarrow B})$, and let ω_1 be the negotiation outcome proposed by agent B obtained from $H^t_{A\leftrightarrow B}$.
2: **if** $u_A(\omega_1) > l^t_A$ **then**
3:    accept the offer (return true).
4: **else if** $u_A(\omega_{best}) > l^t_A$ **then**
5:    not accept (return false) and notify the NB component to propose the negotiation outcome ω_{best} next time.
6: **else**
7:    not accept (return false).
8: **end if**

3.4. Termination-condition (TC) component

This component is responsible for deciding whether to terminate the negotiation and receive the corresponding reservation value or not. Here we treat the reservation value as an alternative offer from the negotiating agent B with a constant utility ru^0_A. Thus the termination conditions of TC component are similar to the acceptance conditions except that $u_A(\omega_1)$ is replaced with the reservation value ru^0_A. The only difference is that we do not need to check the best negotiation outcome proposed by agent B in the history since the reservation value is unchanged throughout the negotiation session. The detailed mechanism for TC component is shown in Algorithm 3.

Algorithm 3. Termination conditions $TC(H^t_{A\leftrightarrow B}, l^t_A, x^t_{A\rightarrow B}, ru^0_A)$.

if $ru^0_A > l^t_A$ **then**
   accept the offer (return true).
else
   not accept (return false).
end if

4. Experimental evaluation and analysis

In this section, we evaluate the negotiation performance of the *ABiNeS* strategy against the state-of-the-art negotiation agents across a wide range of multi-issue negotiation scenarios in the tournament setting following previous evaluation criteria (i.e., efficiency and robustness) (Baarslag et al., 2010; The Second International Automated Negotiating Agent Competition, 2011; The Third International Automated Negotiating Agent Competition, 2012) under the GENIUS (General Environment for Negotiation with Intelligent multi-purpose Usage Simulation) (Lin et al., 2012) platform. The reason why the *ABiNeS* strategy can negotiate more efficiently compared with other state-of-the-art strategies are also investigated in detail, which hopefully can reveal some insights for the automated negotiation community. GENIUS is a negotiation platform of ANAC (Baarslag et al., 2010, 2013; The Third International Automated Negotiating Agent Competition, 2012) developed for facilitating research on bilateral multi-issue negotiations and allowing different negotiation agents to be evaluated in practical environments. All the requirements described in the negotiation model in Section 2 are supported in GENIUS.

4.1. Experimental settings

4.1.1. Negotiation agents

The *ABiNeS* strategy is evaluated against the state-of-the-art negotiation strategies participated in ANAC 2012 (The Third International Automated Negotiating Agent Competition, 2012). The *ABiNeS* strategy is implemented as *CUHKAgent* and has participated the ANAC'12 competition. In this work we evaluate its performance[4] against the state-of-the-art negotiation strategies which enters the final round of ANAC 2012 (The Third International Automated Negotiating Agent Competition, 2012): *AgentLG*, *OMACAgent* (Chen and Weiss, 2012), *TheNegotiatior Reloaded*, *BRAMAgent2*, *Meta-Agent*, *IAMHaggler2012* (Williams et al. 2012a,b), and *AgentMR*. These negotiation agents are designed by different research groups independently and each agent has no information about other agents' strategies beforehand. Each agent can only learn about other agents' information within a single negotiation encounter and all information learned at previous negotiation encounter will be erased at the beginning of the next encounter by the GENIUS platform. The detailed implementations of all negotiation strategies are available at The Third International Automated Negotiating Agent Competition (2012).

4.1.2. Domains

The negotiation domains are designed by different research groups and are targeted at modeling practical multi-issue negotiation scenarios in uncertain and open environments. The domains are different from each other in terms of the number of issues, opposition (Baarslag et al., 2013), discounting factor, reservation values and so on. Opposition is used for reflecting the competitiveness degree of the negotiation domain between the negotiation parties, which is defined as the minimum distance from all the negotiation outcomes to the point representing complete satisfaction of both negotiation parties (1, 1). During negotiations, each agent can only have access to its own preference over all possible negotiation outcomes and has no information about its opponent's preference.

The basic set of the negotiation scenarios consists of 24 different domains submitted by the past participants in previous ANAC competitions. The discounting factor and reservation value can be set to different values within the range of [0, 1], which may greatly influence the performance of negotiation strategies. Therefore, for each negotiation domain from the basic set, three variants are also generated with different discounting factors and reservation values, which thus results in a total of 72 domains (The Third International Automated Negotiating Agent Competition, 2012). The size and opposition degree (competitiveness degree) of the 24 domains are distributed equally. Specifically, there are 8 small-size domains, 8 medium-size domains and 8 large-size domains, and there are 8 domains with high, medium and low opposition respectively. The details of the domains can be found in Williams et al. (2014) and The Third International Automated Negotiating Agent Competition (2012).

4.1.3. Parameter setting of the ABiNeS strategy

The parameter setting of the *ABiNeS* strategy is given in Table 1, which follows the setting used in the *CUHKAgent* implementation. Note that the parameter β is set to different values depending on the value of the discounting factor δ of the current domain.

[4] In this work, we only analyze the negotiation performance among these top eight negotiation strategies in the final round, but it is worth noticing that our *CUHKAgent* also ranks the first place in the qualifying round against a larger number of negotiation opponents (17 different teams).

Table 1
Parameter setting of the *ABiNeS* strategy implemented as *CUHKAgent*.

Parameters	Description	Values
α	Control the way the acceptance threshold is adjusted	0
λ	The minimum value of non-exploitation point	0.08
β	Control the way the non-exploitation point is adjusted with respect to the discounting factor of the domain	1.2/1.5/1.8
γ	Control the way the non-exploitation point is adjusted with respect to the estimated concessive degree of the opponent	10
ω	Weighting factor adjusting the relative effect of the estimated concessive degree of the opponent on the value of the non-exploitation point	0.1
η	The discounting factor reflecting the decreasing speed of the relative importance of the future negotiation outcomes	1
ϵ	The random exploration rate	0.1

Table 2
Overall tournament scores (average score and variance) of the negotiation agents average over different types of domains.

Agent name	Over all domains	Over all discounted domains	Over all undiscounted domains
ABiNeS (*CUHKAgent*)	*0.626478(0.000003)*	*0.577(0.000001)*	0.725(0.000012)
AgentLG	0.621981(0.000003)	0.574(0.000002)	0.717(0.000015)
OMACAgent	0.618437(0.000002)	0.566(0.000001)	0.724(0.000006)
TheNegotiator Reloaded	0.617247(0.000002)	0.555(0.000002)	*0.742(0.000007)*
BRAMAgent2	0.592966(0.000002)	0.565(0.000003)	0.648(0.000009)
Meta-Agent	0.586421(0.000003)	0.551(0.000005)	0.657(0.000004)
IAMHaggler2012	0.535123(0.000001)	0.530(0.000001)	0.546(0.000003)
AgentMR	0.328302(0.000003)	0.361(0.000005)	0.264(0.0)

If $\delta > 0.75$, then $\beta = 1.8$; if $0.75 > \delta > 0.5$, then $\beta = 1.5$; if $0 < \delta < 0.5$, then $\beta = 1.2$.

4.2. Experimental results under tournament

In this section, we present the evaluation results over all domains, discounted domains and undiscounted domains. For each negotiation domain, each negotiation agent negotiates with all other seven agents (excluding itself) for 10 times on both sides in order to produce statistically significant results. All the results described in this paper are adopted from the official final round results of ANAC 2012 (Results of The Third International Automated Negotiating Agent Competition, 2012). We only highlight the overall results comparison and more detailed results can be found at the ANAC 2012 website (Results of The Third International Automated Negotiating Agent Competition, 2012).

4.2.1. Performance over all domains

The second column in Table 2 shows the overall ranking of the eight negotiation agents in terms of overall received utilities and variances averaged over all the 72 domains. We can see that our agent *ABiNeS* achieves the highest average score of 0.626478, ranking the first place. *AgentLG* ranks the second place with the average score of 0.621981, which is slightly lower than that of the *ABiNeS* agent. Both the *OMACAgent* and *TheNegotiator Reloaded* rank the third place since there is no statistically significant difference between their scores.

4.2.2. Performance over discounted and undiscounted domains

When the discounting effect is considered into the negotiation process, the agents need to make careful trade-off between exploiting and concession to the opponent. For the same domain, the negotiation outcomes when the discounting factor is considered are usually significantly different than those cases with no discounting effects. To this end, we further evaluate the negotiation performance of our *ABiNeS* agent against others under discounted and undiscounted domains.

The third column in Table 2 shows the scores of each negotiation agent averaged over all discounted domains. We can see that our *ABiNeS* agent achieves the highest average utility among all agents. We hypothesize that it is because our *ABiNeS* agent can adapt its negotiation decision more efficiently with respect to different discounting factors in terms of balancing between exploiting and concession to the opponent compared with others. Besides, it is worth noticing that the ranks of the *BRAMAgent2* and *TheNegotiator Reloaded* are reversed compared with the previous results averaged over all domains. This may indicated that *BRAMAgent2* agent can make better adaptation to different discounting factors than *TheNegotiator Reloaded* agent under discounted domains.

Next we consider the case of negotiation over undiscounted domains and the scores of all agents averaged over all undiscounted domains are shown in the last column of Table 2. We can see that there are significant changes among the relative ranking of the agents. *TheNegotiator Reloaded* agent ranks the first place while our *ABiNeS* agent ranks the second place. Under undiscounted domains, it is much easier for agents to make decisions since there is no discounting effect on their actual received utilities as time passes. It is usually expected that the less concessive an agent is, the more utility that it will obtain under undiscounted domains. Accordingly, the overall ranking of the agents under undiscounted domains can largely reflect their relative concession degrees.

4.3. Detailed analysis of ABiNeS strategy

Previous section has shown the negotiation power of *ABiNeS* strategy against other state-of-the-art strategies under different negotiation domains. However, it is still unclear which decision-making component contributes to the superior negotiation efficiency of the *ABiNeS* strategy. Having a better understanding of this question can not only enable us to make further improvement of our strategy, but also make it possible for other researchers to incorporate the decision-making components of our strategy into their negotiation strategy design appropriately.

In this section, we mainly focus on investigating two components inside our strategy as follows, which we believe make the most contributions to the overall negotiation performance of the *ABiNeS* strategy:

1. The adaptive adjustment of the non-exploitation point λ based on the behavior of the negotiating partner.

2. The two-stage way in which the acceptance threshold l_A^t is determined.

During negotiations, on one hand, each rational agent would like to obtain as much payoff as possible by exploiting its negotiation partners; on the other hand, due to the discounting effect and time constraints, it is also necessary to make certain compromise to the negotiation partners to avoid obtaining very low utilities in the end (possibly caused by a break-off or high discounting effect). The concession degree should depend on both the characteristic of the negotiation domain and the type of the partner we are negotiating with. Therefore, one key problem during negotiation is how to balance the trade-off between exploiting and making compromise to the negotiating partner. In *ABiNeS*, the adaptive non-exploitation point λ is introduced which represents the specific time when we should stop exploitations on the negotiating partner. The value of λ is adaptively adjusted based on the estimated behavior of the negotiating partner (i.e., its concession degree).

We hypothesize that the adaptive adjustment of λ can help us to better exploit different negotiation opponents, and thus improve the overall utility obtained through negotiation. To evaluate its influence on the overall negotiation performance of *ABiNeS* strategy, we compare the performance of the original *ABiNeS* strategy with the modified *ABiNeS* strategy in which the adaptive adjustment of λ is removed (denoted as *ABiNeS*′). The second column in Table 3 shows the average utilities obtained by both strategies against other seven state-of-the-art strategies in ANAC 2012 in different domains. To make the evaluation results more general, for each domain the simulations are performed over four instances of the domains with different discounting factors ($\delta = 1$, 0.75, 0.5, 0.25). All results are average over 100 times. From the second column in Table 3, we can see that *ABiNeS* strategy can always obtain a higher average utility than *ABiNeS*′ in all negotiation domains. The results thus successfully verify our hypothesis of the usefulness of adaptively adjusting the value of λ based on the estimated behavior of the opponent. For the eight negotiation strategies participated in the tournament, each of them makes concessions during negotiation in various ways. Some of them (e.g., AgentLG) are more aggressive in exploiting the opponents while the rest (e.g., IAMHaggler2012) are less aggressive. In this context, the capability of adaptively predicting the concession degree of different negotiation opponents and making informed move (exploiting more or make more concession) accordingly enables the *ABiNeS* agent to make efficient exploitation against different opponents. Specifically, the *ABiNeS* agent is able to make more exploitation to gain more utility against mild opponents and make concession quickly to avoid low utility against tough opponents. Thus it is expected that *ABiNeS* strategy can obtain higher average utility than *ABiNeS*′ strategy under different negotiation domains, which is confirmed by the simulation results shown in the second column in Table 3.

Table 3
Comparison of the average tournament scores of *ABiNeS* vs. *ABiNeS*′ and *ABiNeS* vs. *ABiNeS*″ in different negotiation domains.

Domain name	*ABiNeS* vs. *ABiNeS*′	*ABiNeS* vs. *ABiNeS*″
Flight Booking	0.593165272/0.544166163	0.595633184/0.576213201
SuperMarket	0.531929334/0.527879116	0.549546851/0.526440695
England vs Zimbabwe	0.642272791/0.641575379	0.708691163/0.586091569
Travel	0.624391428/0.620292902	0.625476427/0.5893065
Outfit	0.637956858/0.633111545	0.6371233/0.599609804
Grocery	0.669997176/0.66421635	0.635214787/0.634578829
Phone	0.681115608/0.674884384	0.670055909/0.623579617
Music Collection	0.737541642/0.731199686	0.74688424/0.716320117
Laptop	0.737647733/0.730444302	0.749388893/0.73743424
IS BT Acquisition	0.790048228/0.780533875	0.774180093/0.77663871
Camera	0.672763642/0.658183578	0.67820488/0.6581175
Housekeeping	0.617283604/0.602394404	0.621411774/0.602394404
Energy (Small)	0.501314241/0.485674969	0.483003949/0.471284284
Itex vs Cypress	0.470921795/0.451796433	0.470921795/0.481204633
Energy	0.399831811/0.379083373	0.399831811/0.379083373
Amsterdam Party	0.673646911/0.652526336	0.668931074/0.647760278
Barbecue	0.674166041/0.63453876	0.650610379/0.642122672
Airport Site Selection	0.582970843/0.541472083	0.569461523/0.560341328
Barter	0.361420434/0.314742612	0.465825872/0.463760361
Rental House	0.526327346/0.479128338	0.465322491/0.46809373

Next we evaluate the influence of the second component, the two-stage way in which the acceptance threshold l_A^t is determined, on the negotiation efficiency of the *ABiNeS* strategy. The two-stage determination of the acceptance threshold works as follows: in the first stage ($t \le \lambda$), the *ABiNeS* agent always makes gradual concession to its opponent following certain pattern of behavior but always maintain its acceptance threshold higher than the predetermined value $u^{max}\delta^{1-t}$; in the second stage ($t > \lambda$), it accepts any proposal in which its discounted utility is no less than $u^{max}\delta$. The rationale behind is that we assume that the opponents are always rational in that any positive proposal would be accepted by the opponent at the very last moment of the negotiation, therefore the worst case is that *ABiNeS* agent obtains the utility of $u^{max}\delta$, i.e., proposing u^{max} at the very last moment of the negotiation to its opponent and the proposal is accepted.

We hypothesize that this assumption is essential in guaranteeing that the *ABiNeS* agent would not be exploited significantly by its opponent during negotiation, and it is expected that the minimum utility $u^{max}\delta$ can always be obtained no matter how aggressive the opponent is. To evaluate the influence of this assumption on the negotiation efficiency of the *ABiNeS* strategy, we compare the average utility of the original *ABiNeS* strategy against the state-of-the-art strategies with the modified version in which this assumption is removed (denoted as *ABiNeS*″). The last column in Table 3 shows the expected utilities obtained by both strategies against other seven state-of-the-art strategies in ANAC 2012 over different domains. Similar with the setting of previous part, for each domain we consider four instances of the domains with different discounting factors ($\delta = 1$, 0.75, 0.5, 0.25). All results are average over 100 times. From the last column in Table 3, we can clearly see that *ABiNeS* agent is able to obtain higher average utility than *ABiNeS*″ agent in all negotiation domains, which thus verifies the importance of adopting this assumption in improving the overall negotiation efficiency. Specifically, for the *ABiNeS* agent, as shown in Fig. 1, the agent's acceptance threshold is usually increased gradually after the time λ, while in contrast, the *ABiNeS*″ agent without the second assumption would just accept any proposal whose utility is higher than $u^{max}\delta^{1-\lambda}$ after the time λ. Therefore, we can see that the *ABiNeS* agent is able to gain more utility from negotiations by exploiting the assumption that the opponents are usually rational in that they would accept any positive proposal at the very last moment of the negotiation.

4.4. The empirical game theoretic analysis of robustness

The evaluation criterion adopted in Section 4.2 reflects the negotiation performance of different strategies in terms of average scores achieved within a fixed tournament setting. However, it does not reveal much information about the *robustness* of the negotiation strategies in different negotiation scenarios, since it assumes that each agent's strategy is fixed beforehand. For example, we may be interested to know whether an agent adopting our strategy *ABiNeS* has the incentive to unilaterally deviate to other strategies under a particular negotiation tournament. Would any agent adopting other strategies be willing to switch to our strategy under different tournaments?

To answer the above questions, we adopt the game-theoretic approach to analyze the robustness of our strategy under different negotiation settings. Since there exists an infinite number of possible negotiation strategies that the agents may take, we cannot apply the standard game-theoretic approach to perform such an analysis by explicitly considering all possible strategies. Therefore, in this paper we adopt the tool of empirical game theoretic (EGT) analysis to achieve this goal instead, which is originally developed by Wellman et al. (2005) to analyze the Trading Agent Competition. EGT analysis is a game-theoretic analysis approach based on a set of empirical results. It handles the problem of the existence of infinite possible strategies by assuming that each agent only selects its strategy from a fixed set of strategies and the outcomes for each strategy profile can be determined through empirical simulations. This technique has been successfully applied in addressing questions about robustness of different negotiation strategies in ANAC 2011 (Baarslag et al., 2013). However, previous robustness analysis is based on the principle of best single-agent deviation, where only the agent with the maximum deviation profit among all participating agents can deviate from its current strategy. However, this assumption may not be able to reflect the realistic scenarios, since each participating agent usually represents different parties and makes decision independently. Each participating agent usually has its own freedom to choose whether to deviate from its current strategy or not. To this end, we propose adopting the principle of *single-agent best deviation* to perform the robustness analysis based on EGT analysis, where each agent is allowed to deviate from its current strategy in a rational way. Therefore, each participating agent has equal opportunity to choose whether it would deviate from its current strategy or not. We apply the EGT analysis to the bilateral negotiation setting as follows.

In our EGT analysis, we consider a fixed set $\mathcal{S}=\{\mathbf{C},\mathbf{L},\mathbf{O},\mathbf{R},\mathbf{B},\mathbf{M},\mathbf{I},\mathbf{A}\}$[5] of negotiation strategies consisting of the top eight strategies from this year's ANAC competition. Different from the setting of ANAC, each agent is free to select any strategy from this set as its negotiation strategy. For each bilateral negotiation, the corresponding payoff received for each participating agent is determined as its average payoff over all possible domains against its opponent, which can be obtained through empirical simulations. The detailed payoff matrix for all possible bilateral negotiations is given in Table 4. Note that for each strategy profile, only the row player's payoff is given since the game itself is agent-symmetric. Based on the bilateral negotiation outcomes in Table 4, it is easy for us to construct the negotiation outcomes (the corresponding payoff profiles) for any possible negotiation tournament among multiple agents. The average payoff of an agent in any given tournament can be determined by averaging its payoff obtained in all bilateral negotiations against all other agents participated in the tournament.

Given a negotiation tournament consisting of n agents, each agent chooses one particular strategy from the set $\mathcal{S}$ of strategies, which jointly constitutes a strategy profile. An agent has the incentive to deviate its current strategy to another one if and only if its payoff after deviation can be statistically significantly improved, provided that all the other agents keep their strategies unchanged. There may exist multiple candidate strategies that an agent has the incentive to deviate to, but here we only consider the best deviation available to that agent in terms of maximizing its deviation benefit. Following previous work (Baarslag et al., 2013), given a pure strategy profile under a negotiation tournament, if no agent has the incentive to unilaterally deviate from its current strategy, then this strategy profile is called a *empirical pure strategy Nash equilibrium*. Another concept for analyzing the stability of the strategy profiles is *best reply cycle* (Yong, 1993), which is a subset of strategy profiles in which, for any strategy profile within this subset, there is no single-agent best deviation path leading to any strategy profile outside the cycle.

Table 4
Payoff matrix for the top eight negotiation strategies in ANAC 2012 average over all domains (for each strategy profile, only the row player's payoff is given since the game is symmetric.) The letters in bold are the abbreviations for each strategy.

Payoff matrix	**C**	**L**	**O**	**R**	**B**	**M**	**I**	**A**
C	0.5956	0.465	0.491	0.669	0.548	0.618	0.832	0.437
L	0.541	0.4212	0.439	0.673	0.462	0.640	0.832	0
O	0.533	0.38	0.4233	0.648	0.433	0.562	0.815	0
R	0.546	0.522	0.502	0.5757	0.509	0.596	0.773	0.425
B	0.523	0.357	0.414	0.657	0.4626	0.648	0.757	0.207
M	0.501	0.486	0.472	0.623	0.484	0.4556	0.76	0.079
I	0.559	0.567	0.55	0.578	0.531	0.592	0.8192	0
A	0.471	0	0	0.615	0.163	0.12	0	0

Both *empirical pure strategy Nash equilibrium* and *best reply cycle* can be considered as two different interpretations of empirical stable sets to evaluate the *stability* of different strategy profiles. The *basin of attraction* of a stable set is defined as the percentage of strategy profiles which can lead to this stable set through a series of single-agent best deviations. Accordingly, the robustness of a strategy is evaluated based on the basin of attraction of the stable set that belongs to Baarslag et al. (2013). Given two strategies s and s', if the basin of attraction of s is larger than that of s', then we say strategy s is more robust than strategy s'.[6]

We evaluate the *robustness* of our strategy *ABiNeS* in the following two negotiation contexts:

- Bilateral negotiations in which each agent is allowed to choose any strategy from the set $\mathcal{S}$.
- Negotiation tournaments following the setting of ANAC 2012 competition, except that the agents are allowed to select the same strategy during the same tournament.

We propose applying model checking techniques (Clarke et al., 1994) to perform EGT analysis by automatically identifying both *empirical pure strategy Nash equilibrium* and *best reply cycle*, and also determining the *basin of attraction* of the above two stable sets. Since the model checking technique is not the focus of the paper, we only show the robustness analysis results in the following part. Interested readers may refer to Hao and Song for details of how to apply model checking to perform robustness analysis.

4.4.1. Bilateral negotiations among eight possible strategies

In the context of bilateral negotiations, there exist a set $\mathcal{P}$ of agents ($|\mathcal{P}|=2$) and a set $\mathcal{S}=\{\mathbf{C},\mathbf{L},\mathbf{O},\mathbf{R},\mathbf{B},\mathbf{M},\mathbf{I},\mathbf{A}\}$ of strategies. Each agent can choose any strategy from the set $\mathcal{S}$ during negotiation. In general there exist a total of $|\mathcal{S}|^{|\mathcal{P}|}=8^2=64$ possible strategy profiles. However, since the bilateral negotiation itself is symmetric (e.g., a negotiation in which agent p_1 uses strategy $\mathbf{C}$ and agent p_2 uses strategy $\mathbf{L}$ is equivalent with the case that their strategies are swapped.), we can simply reduce the

[5] The bold letters are the abbreviations for each strategy as follows: **C**, CUHKAgent; **L**, AgentLG; **O**, OMACAgent; **R**, TheNegotiatorReloaded; **B**, BRAMAgent2; **M**, Meta-agent; **I**, IAMHaggler2012; **A**, AgentMR. These abbreviations will be used in the following descriptions.

[6] Note that the current robustness definition is not complete in that there is no way to compare the relative robustness of strategies within the same empirical stable set.

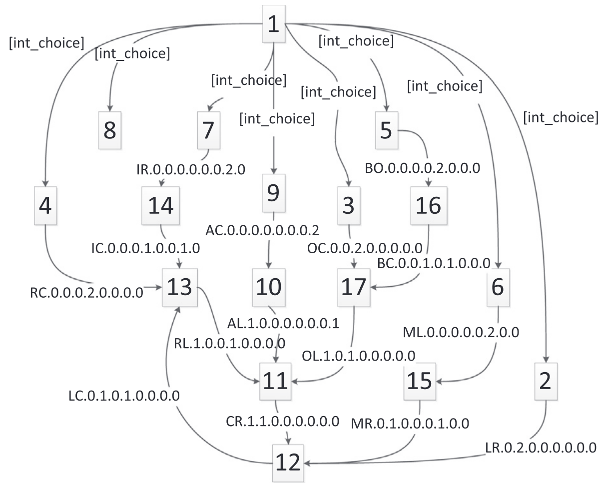

Fig. 3. Deviation analysis with initial state (node 1) consisting of all possible bilateral negotiation settings in which both agents choose the same strategy from the eight possible strategies.

number of strategy profiles from 64 to $\binom{|\mathcal{P}|+|\mathcal{S}|-1}{|\mathcal{S}|-1}=36$. This also indicates that the number of states of the MDP model generated by PAT for performing EGT analysis can be reduced from 64 to 36, which can help in reducing the verification time of PAT.

According to the verification results of PAT (Hao and Song), we find that under bilateral negotiations, there is no *empirical pure strategy Nash equilibrium* and there exists only one *best reply cycle*, i.e., $(L,C)\rightarrow(R,L)\rightarrow(C,R)\rightarrow(L,C)$. Besides, the *basin of attraction* of this cycle is 100%, i.e., for all possible initial strategy profiles, there always exist a single-agent best deviation path which can lead to one of the strategy profiles within this cycle. Our strategy *ABiNeS* is contained in two strategy profiles ((C,R) and (C,L)) in this cycle. This indicates that our strategy *ABiNeS* is very robust against other strategies since it is always possible that the agents will be willing to adopt our strategy *ABiNeS* no matter what their initial strategy is in the long run.

Fig. 3 shows an example of the dynamics of how each agent may change its strategy between different negotiation tournaments under bilateral negotiations among eight strategies, which is generated automatically by the model checker PAT. In Fig. 3, each node (state) is associated with a unique number representing a unique bilateral negotiation setting (strategy profile) except node 1 which, as the initial state, consists of all the possible bilateral negotiation settings in which both agents choose the same strategy. The strategy profile for each state is given by the label of its outgoing transitions. Each transition between any pair of nodes (states) indicates a single-agent best deviation for one particular type of strategy. For example, considering nodes 4 and 13, node 4 represents the bilateral negotiation in which both agents use strategy **R** and node 13 represents the bilateral negotiation in which one agent chooses strategy **C** while the other one chooses strategy **R**, and the transition between them indicates that there exists such a single-agent best deviation from node 4 to node 13. From this graph, we can also easily check that there only exists one *best reply cycle* from node 13 to node 11 to node 12 and back to node 13, and for any initial negotiation settings (strategy profiles) in the initial state (node 1), all of them will finally converge to one node within this cycle through a series of single-agent best deviations.

However, the analysis within bilateral negotiation setting does not give us much information about the robustness of our strategy within a tournament setting involving more than two agents, which will be described in the next section.

4.4.2. Eight-agent negotiation tournaments

In this section, we analyze the robustness of our strategy *ABiNeS* within the context of the negotiation tournament among eight players. We start with a simple case in which each agent is only allowed to choose one strategy from the top four strategies[7] in Section 4.4.2, and then we perform the robustness analysis by taking all top eight strategies into consideration in Section 4.4.2, following the setting of ANAC competition. For both cases, each participating agent negotiates with all the other participates. However, different from the setting in ANAC competition, the agents are free to choose any strategy from the set of strategies available and thus different agents may select the same strategy during the same tournament in our EGT analysis. The setting of ANAC competition can be considered as a specific tournament in which each agent chooses a unique strategy among the eight strategies.

One natural way is to perform robust analysis over all domains, however, this can hide a lot of detailed information due to the averaging effects. Besides, most of the domains are relatively small and thus more easy for the agents to negotiate to get a high utility under the limited negotiation time (3 min). Therefore, we conduct the robustness analysis under the tournament setting over the challenging domain: *Travel* domain, similar to previous work (Williams et al., 2012b). The *Travel* domain is one of the largest and most complex domains in the competition, which thus can better reflect the practical negotiation scenarios which usually involve a large number of possible proposals to consider. In addition, different from Williams et al. (2012b), we set the discounting factor of this domain to the low value of 0.5 instead of 1 (without discounting). Under the setting with high discounting effect, it requires the agents to delicately and adaptively trade off between concession to the opponent (be fear of obtaining lower payoff due to large discounting effect) and staying tough (hope to get higher payoff by letting the opponent concede first) against different types of opponents. Therefore we believe that this setting can better reflect the behavior differences between different strategies.

Negotiation tournament analysis over top four strategies: In this section, we consider the eight-player negotiation tournament over the set $\mathcal{S}=\{C,L,O,R\}$ of the top four strategies. Similar with the analysis in bilateral negotiation, the total number of strategy profiles considered can be reduced from $|\mathcal{S}|^{|\mathcal{P}|}=65\,536$ to $\binom{|\mathcal{P}|+|\mathcal{S}|-1}{|\mathcal{S}|-1}=165$ considering the symmetry of the negotiation.

The robustness analysis is performed using the model checker PAT (Sun et al., 2009). Based on the verification from PAT (Hao and Song), we observe that there only exists one *empirical pure strategy Nash equilibrium*, (C,C,C,C,C,C,C,C), in which all agents adopt our strategy *ABiNeS*, and also for all non-equilibrium tournaments, there always exists a single-agent best deviation path leading to this equilibrium. This result indicates that our strategy is very robust under the setting of negotiation tournament, even though the agents' average payoff under this tournament is lower compared with some other tournaments (e.g., all agents adopting the strategy *IAMHaggler2012*[8]). Similar phenomena (the inefficiency of Nash equilibrium from the social perspective) are commonly observed in non-cooperative game theory. For example, in the prisoner's dilemma game, mutual defection is the only pure strategy Nash equilibrium, but there exists another Pareto-optimal outcome of mutual cooperation under which all agents' payoffs can be significantly increased.

[7] The reason that we choose the top-four strategies instead of the top-three is that both *OMACagent* and *TheNegotiatorReloaded* rank the third place.

[8] This strategy wins the *most social agent* award in ANAC 2012 since it achieves the highest social payoff (the sum of its own and its opponent's payoffs).

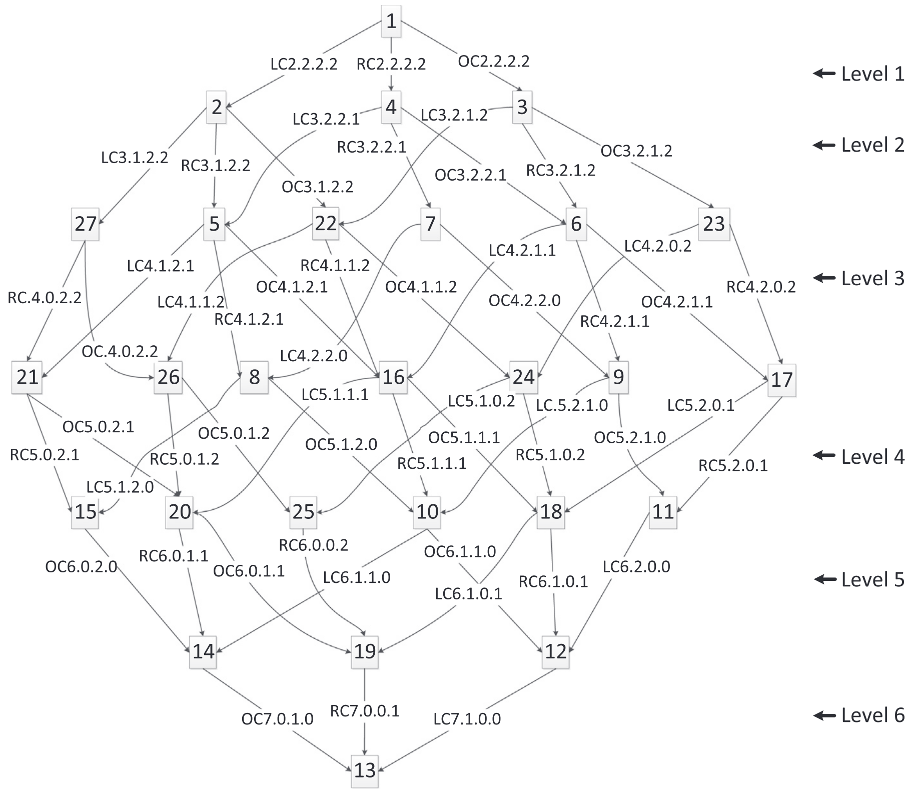

Fig. 4. Deviation analysis with initial state (node 1) in which each strategy is chosen by two agents.

We give a specific illustration of how the agents adjust their strategy choices between different tournaments under the EGT analysis in Fig. 4. This deviation analysis graph is generated automatically using the model checker PAT's simulation tool, which can be interpreted in the same way as Fig. 3. From Fig. 4, we can clearly observe the overall trend that all agents choosing strategies different from C have the incentive to deviate from their current strategies to strategy C. This deviation analysis graph can be viewed as a graph with six levels, and the transitions between every adjacent levels i and j correspond to single-agent best deviations from states in level i to states in level j. At the last level, all agents originally choosing strategies different from C have switched to strategy C, thus converging to the terminating node 13 which corresponds to the pure strategy Nash equilibrium (C, C, C, C, C, C, C, C). Lastly, it is worth noticing that in Fig. 4, some nodes (e.g., node 3) have multiple outgoing transitions, indicating that there exist multiple agents who have the incentives to deviate from their current negotiation strategies. However, previous robustness analysis based on the best single-agent deviation principle fails to model this kind of concurrent deviations.

Negotiation tournament analysis over top-eight strategies: In this section, we turn to the analysis of the complete eight-player tournament over the top eight strategies following the setting of ANAC 2012. In this setting, the total number of strategy profiles considered can be reduced from $|\mathcal{S}|^{|\mathcal{P}|} = 16\,777\,216$ to $\binom{|\mathcal{P}|+|\mathcal{S}|-1}{|\mathcal{S}|-1} = 6435$ due to symmetry of the negotiation. According to the verification results from PAT (Hao and Song), a similar result as that in the previous case in Section 4.4.2 can be observed. There exists only one *empirical pure strategy Nash equilibrium*, (C, C, C, C, C, C, C, C), where all agents adopt our strategy *ABiNeS*, and also the *basin of attraction* of this equilibrium is 100%. This indicates that our strategy *ABiNeS* is also robust under the full negotiation tournament setting over top eight strategies. Similar deviation trends as the case in Section 4.4.2 can also be found here: for any initial tournament, those agents choosing strategies other than strategy C have the incentive to switch to strategy C to maximally increase their individual average payoffs, and thus after a series of single-agent best deviations, finally the strategy profile will converge to the pure strategy Nash equilibrium (C, C, C, C, C, C, C, C).

5. Related work

Until now great effects have been devoted to develop efficient negotiation strategies for automated negotiating agents in the literature. The most typical approach in the literature is concession-based strategy such as ABMP strategy (Jonker et al., 2006). The ABMP strategy agent decides on the next move based on its own utility space only and makes concession to its negotiating partners according to certain concession pattern. However, since the dynamics and influences of the negotiating partners are not taken into consideration, the weakness of this type of strategies is that it is difficult to reach those mutually beneficial outcomes and thus is inefficient in complex negotiation scenarios.

To overcome the aforementioned limitation, various techniques have been proposed to model certain aspects of the negotiation scenario to improve the efficiency of the negotiation outcome such as the opponent's preference profile and the knowledge of the negotiation domain (Faratin et al., 2003; Saha et al., 2005; Hindriks and Tykhonov, 2008; Baarslag et al., 2011; Jakub and Ryszard, 2006; Chen and Weiss, 2012; Williams et al., 2012a; Hindriks et al., 2006; Fatima et al., 2007; Robu et al., 2005; Fujita et al., 2012). Since our work is within the negotiation setting of ANAC competition which focuses on the independent multi-issue negotiation, we only introduce representative works under the settings of the independent multi-issue negotiation in detail. However, it is worth noting that designing effective negotiation strategy and protocol under interdependent multi-issue negotiation setting is also a very active research in the automated negotiation area. Until now, a large number of work has been done in this direction to address the challenges of accurate prediction of the opponent's nonlinear utility function (e.g., the scalability issue with the increasing of the number of negotiating items) (Hindriks et al., 2006; Fatima et al., 2007; Robu et al., 2005; Fujita et al., 2012).

Saha et al. (2005) propose a learning mechanism using Chebychev's polynomials to approximately model the negotiating opponent's decision function in the context of repeated two-player negotiations. They prove that their algorithm is guaranteed to converge to the actual probability function of the negotiating partner under infinite sampling. Experiments also show that the agent using their learning mechanism can outperform other simple learning mechanisms and also be robust to noisy data. However, in their approach, it is assumed that the agents negotiate over one indivisible item (price) only, thus it is not applicable to more general multi-issue negotiation scenarios.

Hindriks and Tykhonov (2008) propose a Bayesian learning based technique to model the negotiating opponent's private preference in the context of bilateral multi-issue negotiations. To make the learning of the opponent's preference feasible, they made two major assumptions: assumption on the structure of the opponent's utility function and the rational assumption of the opponent's tactic. Specifically, they assume that the utility function modeling the opponent's preference is linearly additive and also the evaluation function for each issue can be modeled by means of three types of functions (downhill shape, uphill shape and triangular shape). Besides, they assume that the opponent makes its offer based on a specific linearly concessive function. They evaluate this technique on several negotiation domains and show that it can improve the efficiency of the bidding process by incorporating this preference prediction technique into a negotiation strategy. One notable application of this modeling technique is that it is integrated into a negotiation strategy called *TheNegotiator* (Baarslag et al., 2013) which participated in ANAC 2011 (Baarslag et al., 2013).

Jakub and Ryszard (2006) propose a negotiation strategy based on the predictions of the negotiating partner's future behaviors using difference method. Based on the predictive move of the negotiation opponent, the negotiation can be modeled as multi-stage control process and the task of determining the optimal next-step offer is equivalent to the problem of determining the sequence of optimal control. Simulation results show that the agents using their mechanism can greatly outperform the classical approaches in terms of utilities. However, their mechanism is only applicable in the single-item negotiation scenario. Besides, they put strong assumption on the opponent's behaviors in that the negotiation opponent's strategy is the combination of time-dependent and behavior-dependent tactics, and thus may not work well against other opponents with more complex adaptive strategies.

Chen and Weiss (2012) introduce a novel negotiation strategy OMAC (opponent modeling and adaptive concession), which consists of two major aspects: efficient opponent modeling and adaptive concession making. Different from previous work which directly models the opponent's strategy or its preferences, they propose predicting the OMAC agent's utilities of its opponent's future counter-offers using wavelet decomposition and cubic smoothing spline techniques. Based on the above prediction of the future utilities that the opponent is willing to offer, the OMAC agent adaptively adjusts its concession rate. They evaluate this strategy against the top five negotiation strategies participating in the ANAC 2011 (Baarslag et al., 2013) across a variety of domains under the Genius platform (Lin et al., 2012), and simulation results show that the OMAC agent can outperform the top five agents from ANAC'11. Another notable application of this strategy is that it is implemented as the OMACAgent that has participated in the ANAC 2012 and ranked the third place (The Third International Automated Negotiating Agent Competition, 2012).

6. Conclusion and future work

In this paper, we propose an adaptive negotiation strategy *ABiNeS* for automated agents to negotiate in bilateral multi-issue negotiation scenarios. We introduce the concept of non-exploitation point λ to adaptively adjust the *ABiNeS* agent's concession degree to its negotiating opponent, and propose a reinforcement-learning based approach to determine the optimal proposal for the negotiating partner to maximize the possibility that the offer will be accepted by the opponent.

The performance of the *ABiNeS* strategy is evaluated using two different measures: *efficiency* (average payoff) within a single negotiation tournament and *robustness* which is determined by the size of the basin of attraction of those strategy profiles that our strategy belongs to under different negotiation tournament settings. Our strategy *ABiNeS* is shown to be very efficient against the state-of-the-art strategies from ANAC 2012 and can obtain the highest average payoff over a large number of negotiation domains. Detailed analysis of the *ABiNeS* strategy in terms of the influences of its two major decision components on the negotiation efficiency is also provided, which gives us valuable insights of why it can win the championship in ANAC 2012. Last but not the least, we propose applying model checking techniques to perform EGT analysis to determine the robustness of the strategies, and the *ABiNeS* strategy is found to be very robust in both bilateral negotiations and negotiation tournaments among eight players following the setting of ANAC competition.

We believe that the performance of the *ABiNeS* strategy can be further improved in a number of ways. As future work, one worthwhile direction is to further refine the estimation of the negotiating partner's concessive degree to further exploit the negotiating partner, by taking into consideration the magnitude of the utility that the negotiating partner proposes. Besides, it is worthwhile to further investigate the negotiation performance in ANAC 2012 and integrate the current existing negotiating techniques (i.e., preference modeling and strategy prediction) into the *ABiNeS* strategy under different combinations in different negotiation scenarios.

Acknowledgments

This work is supported by National Natural Science Foundation of China (NSFC) 61170077, SZ-HK Innovation Circle project ZYB200907060012A, NSF GD: 10351806001000000, GDS&T 2012B091100198, S&T project of SZ JC200903120046A, and CUHK Direct Grant for Research (Project ID 4055024).

References

Baarslag, T., Hindriks, K., Jonker, C., Kraus, S., Lin, R., 2010. The first automated negotiating agents competition (anac 2010). New Trends in Agent-Based Complex Automated Negotiations, 113–135.

Baarslag T, Fujita K, Gerding E H, et al. Evaluating practical negotiating agents: Results and analysis of the 2011 international competition[J]. Artificial Intelligence, 2013, 198: 73-103.

Baarslag, T., Hindriks, K., Jonker, C., 2011. Acceptance conditions in automated negotiation. In: ACAN'11.

Chen, S.Q., Weiss, G., 2012. An efficient and adaptive approach to negotiation in complex environments. In: Proceedings of the Twentieth European Conference on Artificial Intelligence, pp. 228–233.

Clarke, E.M., Grumberg, O., Long, D.E., 1994. Model checking and abstraction. ACM Trans. Program. Lang. Syst. 16 (5), 1512–1542.

Coehoorn, R.M., Jennings, N.R., 2004. Learning an opponent's preferences to make effective multi-issue negotiation trade-offs. In: ICEC'04. ACM Press, pp. 59–68.

Faratin, P., Sierra, C., Jennings, N.R., 2003. Using similarity criteria to make negotiation trade-offs. Artif. Intell. 142 (2), 205–237.
Fatima, S.S., Wooldridge, M., Jennings, N.R., 2007. Approximate and online multi-issue negotiation. In: Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems, p. 156.
Fujita, K., Ito, T., Klein, M., 2012. An approach to scalable multi-issue negotiation: decomposing the contract space. Comput. Intell. 2, 399–406.
Hao, J.Y., Song, S.Z. Online Technical Report. ⟨http://www.comp.nus.edu.sg/~pat/eaai/⟩.
Hindriks, K., Tykhonov, D., 2008. Opponent modeling in auomated multi-issue negotiation using Bayesian learning. In: AAMAS'08, pp. 331–338.
Hindriks, K., Jonker, C.M., Tykhonov, D., 2006. Eliminating interdependencies between issues for multi-issue negotiation. In: Cooperative Information Agents X, pp. 301–316.
Jakub, B., Ryszard, K., 2006. Predicting partner's behaviour in agent negotiation. In: AAMAS '06, pp. 355–361.
Jonker, C., Robu, V., Treur, J., 2006. An agent architecture for multi-attribute negotiation using incomplete preference information. AAMAS 15 (2), 221–252.
Kersten, G.E., Noronha, S.J., 1999. Www-based negotiation support: design, implementation, and use. Decis. Support Syst. 25 (2), 135–154.
Lin R, Kraus S, Baarslag T, et al. Genius: An integrated environment for supporting the design of generic automated negotiators. Computational Intelligence, 2012.
Osborne, M.J., Rubinstein, A., 1994. A Course in Game Theory. MIT Press, Cambridge.
Results of The Third International Automated Negotiating Agent Competition (ANAC 2012). ⟨http://anac2012.ecs.soton.ac.uk/results/final/⟩.
Robu, V., Somefun, D., La Poutré, J.A., 2005. Modeling complex multi-issue negotiations using utility graphs. In: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 280–287.
Saha, S., Biswas, A., Sen, S., 2005. Modeling opponent decision in repeated one-shot negotiations. In: AAMAS'05, pp. 397–403.
Sun J, Liu Y, Dong J S, et al. Integrating specification and programs for system modeling and verification[C]//Theoretical Aspects of Software Engineering, 2009. TASE 2009. Third IEEE International Symposium on. IEEE, 2009: 127-135.
The Second International Automated Negotiating Agent Competition (ANAC 2011). ⟨http://www.itolab.nitech.ac.jp/ANAC2011/ ⟩.
The Third International Automated Negotiating Agent Competition (ANAC 2012). ⟨http://anac2012.ecs.soton.ac.uk/⟩.
Wellman, J.E.M.P., Singh, S., Vorbeychik, Y., Soni, V., 2005. Strategic interactions in a supply chain game. Comput. Intell. 21 (1), 1–26.
Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R., 2012a. Negotiating concurrently with unknown opponents in complex, real-time domains. In: ECAI'12, pp. 834–839.
Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R., 2012b. Using gaussian processes to optimise concession in complex negotiations against unknown opponents. In: IJCAI'12, pp. 432–438.
Williams, C.R., Robu, V., Gerding, E.H., Jennings, N.R., 2014. An overview of the results and insights from the third automated negotiating agents competition (anac2012). Novel Insights in Agent-based Complex Automated Negotiation, Studies in Computational Intelligence 535, 151–162.
Yong, H., 1993. The evolution of conventions. Econometrica 61 (1), 57–84.
Zeng, D., Sycara, K., 1996. Bayesian learning in negotiation. In: AAAI Symposium on Adaptation, Co-evolution and Learning in Multiagent Systems, pp. 99–104.
Zeng, D., Sycara, K., 1998. Bayesian learning in negotiation. Int. J. Human Comput. Syst. 48, 125–141.