Value Model Agent: A Novel Preference Profiler for Negotiation with Agents

Asaf Frieder and Gal Miller

Abstract. As multi-agent systems become more complicated, situations arise were agents have different goals. Accomplishing their goals may require trading resources or compromising on the state of the system. In these situations cooperation can improve the utilities of all parties, but some agreements are preferred to others. Humans have reached these agreements for thousands of years through an art known as negotiation. In this paper we describe an agent that was developed for the ANAC2011 bilateral negotiation competition. Our main contribution is a novel approach to modeling the preference profile of the other agent. This preference profile is then used to improve exploration of the bid space and approximate the opponent's concessions.

1 Introduction

For thousands of years humans have resolved conflicts using negotiation. From children deciding what to play to business deals and peace agreements, negotiation is an integral part of human interaction. In some multi-agent systems much like in a community, agents have different goals or preferences. In these cases the agents may conflict over their need of resources or their desired world state. The agents may interfere with each others' plans by withholding necessary resources or changing the world state. It can therefore be beneficial for agents to agree on a common coordinated plan that is more efficient for both agents even if not optimal for either. However, in most negotiation domains there are multiple possible plans (bids) that can be chosen and the agents have different preferences. Thus, an agent needs negotiation skills to ensure its own preferences will be prioritized in the final agreement.

The *ValueModelAgent* was designed to negotiate with other agents as part of the ANAC2011 negotiation competition. Our primary effort was focused on developing a heuristic for reconstructing the profile of the other agent's preferences described in Section 2. An accurate model of the opponent's preferences has many possible

T. Ito et al. (Eds.): Complex Automated Negotiations, SCI 435, pp. 199–203.
springerlink.com

applications in negotiation. The model can be used to increase the likelihood of the opponent accepting bids, by choosing the bids our opponent prefers. The preference profile can also be used to calculate reservation values based on measurements of fairness such as the nash and kalai points. The opponent's concession rate can also be calculated using a preference profile. Using the concession rate, the agent can model the opponent's behavior and estimate the future agreement under different bidding strategies.

We farther describe our bidding strategy that uses the profile to select bids as well as approximate our opponent's concession in Section 3. Our bidding strategy is rule based and takes both our opponent's concessions and the timeline into account.

2 Modeling Preferences

The ValueModel of our agent attempts to extrapolate the opponent's preference profile from the order of our opponent's bids. However, our heuristic requires utility approximations for our opponent's bids. We therefore must generate initial approximations for these bids based on their order. The *a priori* approximations for the opponent's bids are generated based on the assumptions that the opponent makes bids in the order of their preference profile and that the preference profiles of both parties are roughly symmetric. Under these assumptions, the *a priori* approximation for the i-th bid proposed by our opponent is approximate to the utility of the i-th highest bid in our preference profile. While the inaccuracies in these assumptions may cause noise in the *a priori* approximations, they serve only as a starting point for the heuristic we will now describe.

In the genius platform a bid B is a set of values chosen for each of N issues ($B = v_1..v_N$). Recall that a negotiation domain has several issues, each with a different weight w_I for our opponent. Each issue has several values, with different scores. The ratio between a value's score and the maximal score in the issue, is the portion of the issue's weight that our opponent will gain if that value were to be chosen. We define value utility loss (*VUL*) as the maximal difference between our opponent's utility for two bids if one contains the value and the other can only differ in the value for that issue ($I(v)$) (see Eq. 1). We farther define the bid utility loss (*BUL*) as the sum of the utility lost by all values of the bid (or concession).

$$VUL(v) = w_{I(v)} \cdot \left(1 - \frac{eval(v)}{\max\limits_{v' \in I(v)} eval(v')}\right) \tag{1}$$

In the value model, each value of each issue has three properties, the *VUL* approximation, the reliability (*rel*) the agent associates to that approximation and a weighted deviation (*sd*) measurement. When we receive the opponent's first bid, we assume the values in that bid are optimal ($\forall i, VUL(v_i) = 0$), with very high reliability (0.9). We also initialize all other values with high utility loss ($VUL = 1/N$), a very low reliability (0.02) and high deviation ($1/N$). We will now describe how we update these properties when we receive new bids.

Let $BUL(B)$ be the *a priori* approximation for the utility lost by our opponent's new bid. For each value v_i in bid B our model already has an approximation $VUL(v_i)$, and the sum of these approximation is an alternative approximation for $BUL(B)$. Our heuristic needs to be updated based on the difference between the approximations. Unfortunately the model's approximation for BUL is derived from several VUL variables and we need to decide which values need to change. We propose that the values with the lowest reliability and highest deviation are most likely responsible for the difference. Equation 2 demonstrates the calculation of the number of units the set of values should move. We then generate new approximations for each value, using Equation 3 such that the sum of these approximations is the a priori BUL approximation. In addition we generate a reliability measurement for these approximations using a norm formula on the (inverse) original reliabilities of the bid's values as demonstrated in Equation 4.

$$units = \frac{BUL(B) - \sum_{i=0} VUL(v_i)}{\sum_{i=0}^{n} \frac{1}{\sqrt{rel(v_i)}} \cdot sd(v_i)} \tag{2}$$

$$VUL^B(v_i) = VUL(v_i) + units \cdot sd(v_i) \cdot \frac{1}{\sqrt{rel(v_i)}} \tag{3}$$

$$rel^B(v_i) = \sqrt{\frac{1}{n} \cdot \sum_{i=0}^{n} \frac{1}{rel(v_i)^2}} \tag{4}$$

We are now left with two sets of VUL approximations and reliabilities for each value: the model's prior approximations and the ones generated for the new bid. Our objective is to merge these approximations, taking their reliabilities into account such that reliable values will be updated based on temporal difference learning using a learning rate $\alpha = 0.1$ and unreliable values will be updated faster. Therefore, the new approximation takes up to α from the weight (rel) of the original approximation and also takes a portion of the weight not claimed by the original approximation ($1 - rel$). In both cases the portion of the weight received by the new approximation depends on its reliability (rel^B). The resulting weight formulas are displayed in Equation 5 and 6, respectably.

$$p = rel(v_i)(1 - \alpha \cdot rel^B(v_i)) \tag{5}$$

$$p^B = rel^B(v_i)(1 - rel(v_i) + \alpha rel(v_i)) \tag{6}$$

We now update the three properties of the value. We first set the value's VUL approximation to be the weighted average of the two approximations (see Eq 7). Then we estimate a weighted standard deviation based on the distance of VUL^B and the original distribution (based on VUL and sd) from the new approximation VUL'.

Finally, we update the reliability of the approximation based on the weights and how accurately the approximation represents VUL and VUL^B (See Eq. 8).

$$VUL'(v_i) = \frac{p \cdot VUL(v_i) + p^B \cdot VUL^B(v_i)}{p + p^B} \tag{7}$$

$$rel' = p\left(1 - \frac{|VUL(v_i) - VUL'(v_i)|}{2sd'(v_i)}\right) + p^B\left(1 - \frac{|VUL^B(v_i) - VUL'(v_i)|}{2sd'(v_i)}\right) \tag{8}$$

3 Bidding Strategy

At any point in the negotiation the agent has a minimal utility threshold and permits itself to propose any bid above that threshold. As long as there are unproposed bids above the threshold in 80% of the bids we propose the unproposed bid with the highest utility for the opponent according to the value model. In 20% of the bids we pick bids with values that were not contained in recent bids. In cases all bids above the threshold were already proposed, we randomly pick a bid from the bids with the highest 25% utilities for the opponent.

Our threshold strategy depends on several factors[1] including the elapsed time. In the beginning of the negotiation we set our threshold to be 0.98. Until 80% of the time elapses we employ a very strict concession strategy. We only lower the threshold if our opponent conceded 50% more than us based on the average (BUL) of the last 5 new bids our opponent proposed. The second requirement is that our opponent has stopped conceding recently. If both requirements are met we gradually lower the threshold by up to 0.02 over a 5% interval of the timeline.

After 80% of the time has elapsed we measure the utility difference between our current threshold and the opponent's highest bid. For the next 10% of the timeline we gradually lower our threshold by up to 50% of this difference but no lower 0.7. than if our . We also restrict our concession to be no more than 0.05 beyond our opponent's concession. We also require that our opponent has stopped making significant concessions.

After 90% of the time elapsed, we try to use scare tactics to convince our opponent to accept our bids by masquerading as ultimatums. We intentionally sleep three times, for half the remaining time, followed by the approximated best bid for our opponent above our threshold. After each of these bids we lower our threshold and resume exploring bids for a short amount of time. If the opponent did not accept our first proposal, we lower our threshold as low as 0.65 and after the second we lower it to 0.6 [2]. If our opponent rejected the third ultimatum, and our opponent's best bid is above 0.55 we propose it. If our opponent's best bid is bellow 0.55, we resume exploring bids above the threshold.

[1] Due to space restrictions we do not mention some of these factors, most notably the discount factor.

[2] In both cases we lower our threshold by only a portion of the utility difference.

4 Conclusions

Our approach is not without faults, and our agent was in eighth place (eighth) in the finals (third in the qualifying rounds). While reconstructing a preference profile may have interesting applications we did not leverage our model much in our bidding strategy. Identifying the best bids for the opponent may increase the chance of the ultimatums being accepted, but its primary use in the exploration is futile since in most cases there is enough time to propose all bids. Using the approach to approximate our opponent's concession has more potential, but was not fully utilized. The effort that was invested in profiling our opponent's preferences would have been much better spent in profiling our opponent's strategy and adjusting our strategy to our opponent's type. For this end we could have measured his concessions using the existing bayesian model, or even the symmetry assumption.

Finally, making serious concessions after only 80% of the time passed is a mistake. Agents can send many proposals in a very short amount of time, and agents that employ hard bargaining strategies will likely wait longer and take advantage of our "early" concessions. An agent that waits longer, can then utilize its preference model to rapidly send the best bid for the opponent under different thresholds.