Navajo Agent: Minimizing Distance to Nash in Bilateral Negotiations

Nikolaos Chrysovalantis Tsagkopoulos
University of Southampton
ncdt1u18@soton.ac.uk

Konstantinos Mertzanis
University of Southampton
km2n18@soton.ac.uk

Dimitri Diomaiuta
University of Southampton
dd2u18@soton.ac.uk

ABSTRACT
This report proposes a bilateral negotiation strategy for the class tournament of the course COMP6203 based on the specifications of ANAC 2018 competition. Our agent proposes the best bid for both parties according to its own viewpoint while tries to maintain high individual utility.

1. INTRODUCTION
Navajo Agent's strategy is based on the guidelines ANAC 2018 competition initially proposed. It has been developed to handle and estimate its own and opponent's preferences through uncertainty profiles, placing bids that maximize social welfare[1] and accepting bids similar the already offered. This strategy can be decoupled in bidding, opponent modeling and accepting phase as BOA architecture suggests [1].

2. STRATEGY
2.1 Opponent modeling
Agents have to estimate the preference profile of their opponents in order to configure a bidding strategy. Navajo agent uses the same opponent modeling as proposed by "Jonny black" agent which took part in the ANAC 2015 [2]. The goal is to evaluate the importance of every issue and value based on the received bids. The concept is grounded on two assumptions:

- The most desired values for the opponent appear repeatedly in the received bids.
- Most significant issues are less likely to be changed.

The agent stores the bids every time in a frequency table and after a specific number of bids calculates updates the weights of the issues and the importance of the values. The importance of the options of each issue are ordered as:

$$f_0 = \frac{k - n_0 + 1}{k} \quad (1)$$

Where n_0 rank of option o (the most frequent option has rank 1) and k is the cardinality of the set of options. Because of the fact that frequencies range from $[\frac{1}{k}, 1]$, f_0 can express probabilities. Because of this, the unnormalized weights can be obtained through the Gini-Impurity score [3, 4].

$$\widehat{W_\iota} = \sum_{o \in O_i} \frac{f_0^2}{t^2} \quad (2)$$

Then, weights can be normalized easily by:

$$w_i = \frac{\widehat{W_\iota}}{\sum_{j \in I} \widehat{W_j}} \quad (3)$$

The aforementioned strategy gives a more complete opponent preference profile as more bids are being received, that is as the time passes. We found it more efficient to update the weights according to the size domain; for small domains weights update in every bid and for larger domains less and less often, but it was shown that even in bigger domains the agent could handle updating the weights every 10 bids.

2.2 Bidding strategy
Navajo aims to an agreement that maximizes social welfare while securing a high utility. Its bidding strategy is highly inspired by Agent H [5]. To find Nash bargaining solution, it has to create a good opponent model first. In order to achieve that and avoid making a bad bid starts with setting a utility value $U_{TH} = 1$ which declines regarding time and does not offers bids with own utility below this value.

The time conceding value of utility U_{TH} derives from a time dependent tactic, with exponential decay:

$$U_{TH} = 1 - \left(\frac{\min(t, T)}{T}\right)^{\frac{1}{\psi}} \quad (4)$$

```
for t in time:
    utility_th ¬ 1-t**2
    if lastReceivedOffer == sameSendingOffer:
        Offer(BidWithRange([utility_th,1]))
    else:
        best_socialWelfare
                ¬ own_utility(bid[0])*opp_utility(bid[0])
        best_bid ¬ bid[0]
        for b in bids([utility_th,1]):
            if own_utility(b)*opp_utility(b) > best_socialWelfare:
                best_socialWelfare ¬ own_utility(b)*opp_utility(b)
                best_bid ¬b
```

Figure 1. Pseudocode of the bidding strategy.

Where t current time, T maximum negotiation time and ψ a constant [6]. The agent follows a Boulware tactic ($\psi = 0.3$) in

[1] The absence of discount value turns Nash bargaining solution to be equivalent with the bid which maximizes the social welfare in bilateral negotiations

order to secure more time creating the opponent preference profile. During the negotiation, Navajo searches all the possible bids with own utility in the range $[U_{TH}, 1]$ and offers the bid where:

$$bid = argmax_{b \in B}(U_{own}(b) \times U_{opp}(b)) \qquad (5)$$

Assuming than the agent has some knowledge of the preference profile of its opponent, which stops to update from a point and on, as the U_{TH} decays it still offers the best bid it found. In other words, tries to persuade its opponent that this bid is the best for both.

If the agent detects that its opponent is very persistent (hard headed tactic) it just concedes with time as U_{TH} decays, offering random bids with utility $[U_{TH}, 1]$ to avoid not reaching an agreement versus behavioral based agents [7]. Walking away, or unsuccessful negotiation do not give any points, so we eliminate such scenarios.

2.3 Accepting strategy

The accepting strategy is very simple, it accepts an offered bid if it satisfies one of two conditions: a) its closer to Nash bargaining solution than the bid we last offered. It is clear, since the received bid is evaluated by Navajo, that opponent and own preference profile estimation has a huge impact in the outcome of the negotiation as not only can lead in offering bad bids but also not accepting a good one; b) the received bid gives to our agent more utility than the other agent gets.

```
if (social_welfare(ReceivedBid) > social_welfare(OfferedBid)) -delta
    or (own_utility(ReceivedBid) > opp_utility(ReceivedBid)):
    Accept(ReceivedBid)
```

Figure 2. Pseudocode for accepting strategy.

A constant δ proportional to the domain size introduced, to make sure that our agent will accept very similar bids.

3. DEALING WITH UNCERTAINTY

The initial plan for defining our agent's preference profile it was to use "Jonny black's" concept again. Specifically, the only information for our preference profile is the bid ranking. The bid ranking consists of a sequence of ordered bids according utility. A successful frequency table is created by a pre-built heuristic function which gives 1 point to all options of the issues occurred in the lowest ranked bid, 2 points if they occurred in the second lowest ranked bid, 3 points etc. Then, working the weights on this frequency table could result in enough good estimation of our profile. Due to technical reasons of not having access to this table the method was not implemented, and the pre-built weight estimation function was used.

Nevertheless, the prebuilt estimation function works well in big domains but lacks accuracy in small domains.

4. RESULTS

As metrics we calculated the average utility and the average distance to Nash of our agent, comparing the average performance of all agents.

Table 1. Agreement rate per domain.

	small	medium	large
Navajo	96.39%	59.36%	71.98%
Total Avg.	95.49%	72.25%	82.07%

Table 2. Average utility per domain.

	small	medium	large
Navajo Agent	0.946	0.661	0.791
Total Avg.	0.685	0.648	0.654

Table 3. Average Distance to Nash per domain.

	small	medium	large
Navajo Agent	0.070	0.510	0.334
Total Avg.	0.147	0.763	0.552
Min/Max D.	0.049/0.2	0.343/0.8	0.221/0.6

Table 4. Comparing our agent against the top 5 of the tournaments.

	small		medium		large	
	Utility	Nash	Utility	Nash	Utility	Nash
Agent 1	0.995	0.049	0.728	0.443	0.945	0.298
Agent 17	0.957	0.049	0.599	0.420	0.872	0.226
Agent 4	0.950	0.064	0.578	0.383	0.816	0.226
Agent 38	0.949	0.063	0.714	0.460	0.830	0.234
Agent 2	0.957	0.062	0.760	0.383	0.815	0.290
Navajo	0.946	0.070	0.661	0.510	0.791	0.334

From the agreement rate we expected to estimate how well the "Jonny black" method performed in large domains. We assumed that in large domains less offers where made causing in fewer agreements. Unfortunately, apparently Table 1 do not give such a metric.

While Navajo agent secured a better average distance to Nash than the majority of other agents, the great distance with the low agreement rate is enough to describe the implemented preference estimation strategy insufficient.

5. CONCLUSION

Our suggested method performed good overall in the tournament among other agents. The bidding part was effective, rational and values social welfare more than personal utility.

Agent's personal profile modeling was naïve and could be tuned better if we could overcome the technical problems. However, opponent's profile modeling was good and handy even it could be improved more.

Uncertainty was a great challenge for someone to put hands on, as it is a very recent topic and it can be seen in many virtual assistants as Siri and other services such as Netflix and Spotify. These services do not have full access to the preferences of their users and try their best to recommend them related products.

6. REFERENCES

[1] Baarslag, T., Hindriks, K., Hendrikx, M., Dirkzwager, A. and Jonker, C. Decoupling negotiating agents to explore the space of negotiation strategies. Springer, City, 2014.

[2] Yucel, O., Hoffman, J. and Sen, S. *Jonny Black: A Mediating Approach to Multilateral Negotiations*. Springer International Publishing, City, 2017.

[3] Mohri, M., Rostamizadeh, A. and Talwalkar, A. *Foundations of Machine Learning*. The MIT Press, 2012.

[4] Wikipedia contributors, *Decision tree learning,* *https://en.wikipedia.org/wiki/Decision_tree_learning*

[5] Hayashi, M. and Ito, T. *AgentH*. Springer International Publishing, City, 2017.

[6] Faratin, P., Sierra, C. and Jennings, N. R. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, 24 (1998), 159-182.

[7] Baarslag, T., Hindriks, K. and Jonker, C. A tit for tat negotiation strategy for real-time bilateral negotiations. Springer, City, 2013.