

ANAC 2022 Academic report

Authors: Ofek Tal, Igor Mayor, Or Politi

In this report we will present our bidding agent, which we call the “Greedy Entropy Agent”, or GEA in short.

Bidding Strategy:

As in real life, the main challenge in striking a deal is to maintain a reasonably profitable outcome for ourselves, while meeting the other party’s needs in order to achieve an agreement.

Therefore, the goal of our agent is to behave reasonably against the other party while not losing sight of our own interest. To achieve that, we balance our own needs (a high utility score) with an evaluation of the other agent’s needs (their own utility score).

A common and reasonable assumption in game theory is that the opponent behaves rationally – and thus, we assume that the other agent would only offer bids that satisfy them according to their own standards. Likewise, if we offer a bid, it would only be accepted if we meet the other agent’s criteria.

Based on the assumption above, we can define the following theoretical opponent estimation function:

$$\text{model}(\text{bid}) = \begin{cases} 1, & \text{Other party is estimated to accept the bid} \\ -1, & \text{Other party is estimated to reject the bid} \end{cases}$$

Weighting this model in a heuristic score will allow us to prefer bids that we estimate to be accepted, while penalizing bids that we estimate to be rejected. Since we do not have access to the other agent’s utility function, we’ll have to use a binary estimation of their reaction instead. We will expand on how we implement this model in the “Opponent modeling” section.

For now, assuming that we have such a model, we can formally define a heuristic function. Given the time passed since the negotiation started, and an “agreement weight” which dictates how much attention we assign to the opponent prediction, we suggest the following function:

$$H(\text{bid}, \text{time}) = 0.95 * (1 - \text{time}^{10}) * \text{utility}(\text{bid}) + \text{model}(\text{bid}) * \text{agreement_weight}$$

The first addend takes into account the time pressure and our own utility, while the later addend takes into account the opponent’s expected reaction to the bid. The agreement weight is dynamic and we will expand on how it is calculated in the “Learning method” section.

We randomly sample **500** bids, uniformly distributed across the domain, and select the highest heuristic value found.

Acceptance Strategy:

Using the heuristic function defined above, the acceptance strategy is rather straightforward.

We accept a bid if either:

- The utility score of the bid is higher than 0.7, meaning that this bid is quite valuable for us.
- The heuristic function yields value higher than 0.85. This value was selected after fine-tuning.
- The time progression of the session is above 0.9 and the utility of the bid is higher than 0.6, it is better to settle on a “mid-range” bid than to gain nothing.
- The time progression of the session is above 0.95 and the utility of the bid is higher than 0.4, representing an even softer threshold in case time is pressing.

Opponent modeling:

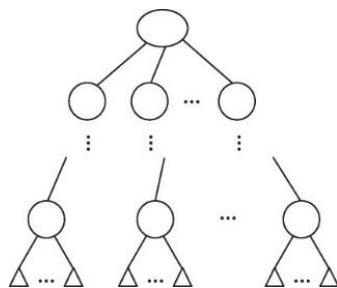
Decision tree:

In practice, we implement the opponent prediction model using a decision tree classifier. A bid is converted to a vector of 0's and 1's, based on the values it contains.

We collect all the past offers of the other agent, and label them **1** since we assume our opponent would not suggest a bid unless they are ready to accept it.

We also collect our own bids and label them **-1**, since the other agent would either reject the bids, meaning that they are not ready to accept them, or they accept and the session ends anyway.

Finally, we train a decision tree classifier on the complete dataset. We limit the decision tree to a maximum depth of 20 to allow for better generalization.



Learning method:

Agreement weight:

Our goal is to dynamically adjust this parameter to distinguish between greedy and reasonable agents.

A reasonable agent is an agent that tries to match our preferences over time, while unreasonable agents are agents that do not seem to strive for a settlement, and behave in a “hardliner” way.

Against reasonable agents, we can offer less valuable bids in hopes to reach an agreement. On the other hands, unreasonable agents would not strive to meet our needs, and therefore we are reluctant to offer and accept bids that would satisfy them.

The agreement weight begins with a value of **0.15**. Then, whenever the other agent offers a bid, we compare it's utility (our utility) against their previous offer's utility (our utility). If the value increases, it means that the agent tries to satisfy us. In this case we multiply the agreement weight by $\frac{1}{0.99} = \sim 1.01$ to assign more importance to the opponent's agreement prediction.

If the utility value decreases, it means the other agent fails to adapt to our needs. In this case, we multiply the agreement weight by **0.99** to achieve the opposite effect.

Overtime, agents that show a relatively consistent trend would receive larger or smaller weights based on their behavior.

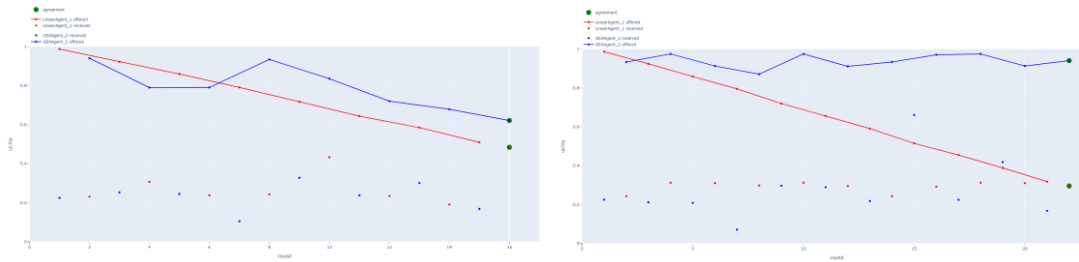
Analysis:

Against boultware agent:



During most of the session, a boultware agent is persistent and would not settle for less-than-ideal bids. Our agent replicates that behavior, weighting the opponent's expected utility less against our own as time progresses. In the last few rounds, we offer a reasonable bid as the balware agent starts to show more altruistic behaviour.

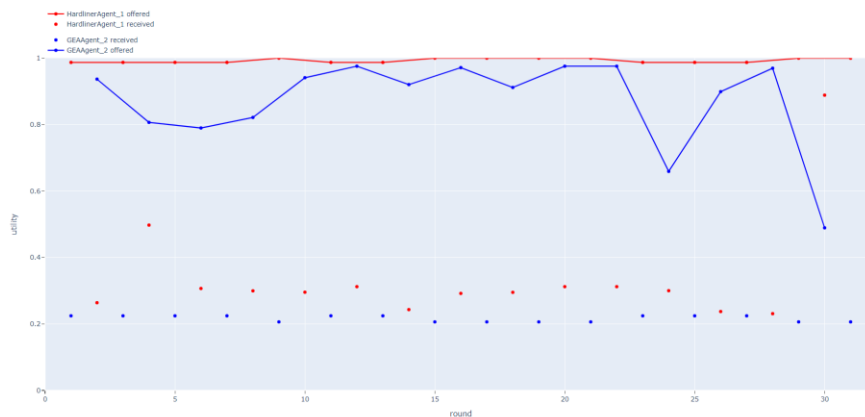
Against linear agent:



(Left graph:) The linear agent teaches our model that it is willing to accept a wider variety of bids, and therefore our model is able to offer bids that are relatively more beneficial to the other agent. A linear agent is an example of a reasonable opponent – we are able to reach an agreement over a bid that is reasonably beneficial for both of us.

(Right graph:) For the sake of comparison, we locked the agreement weight to **0** to see how the learning process affects our model. As expected, without learning the other agent’s preferences our model behaves in a strictly greedy way, offering the bid with the highest utility (our utility) found in the sampling process. In this case, the linear agent agreed to our greedy offer - however against agents who are more competitive that would not be a likely outcome.

Against hardliner agent:



A hardliner agent is an example of an extremely unreasonable opponent. Since this agent is unable to learn our preferences in order to reach a settlement, or offer a wide variety of bids, our agent shows a similar greedy behavior, offering only the highest sampled bids that we estimate to be accepted. In this case, the agents do not reach an agreement before the session ends. It is important to note that during the actual competition, hardliner agents might perform better in practice if a large number of contenders were to concede to their repeating offers. Despite that, we still strive to develop a reasonable agent since we prioritize our academic learning process.