

The Cunning Merchant: A Deceptive Bidding Strategy for the ANL 2026 Deception Challenge

Kaan Gönenli¹, Umut Ardil Murat²

¹Department of Computer Science, ²Department of Computer Science

Student ID: ¹S034351, ²S028641

¹kaan.gonenli@ozu.edu.tr, ²ardil.murat@ozu.edu.tr

CS 451: Introduction to Artificial Intelligence — Spring 2026

Abstract

This report presents the design, implementation, and evaluation of the “Cunning Merchant” negotiation agent for the ANL 2026 Deception Challenge. The agent is evaluated on two criteria: the utility it gains from agreements, called Advantage, and how effectively it conceals its private utility function from the opponent, called Concealing.

Our agent combines six complementary tactics: a strong opening threshold to anchor the negotiation, a Boulware-style concession curve with exponent $e=4.0$, random haggling fluctuations to disrupt simple opponent curve-fitting, preference-inverting bids that mislead the opponent’s frequency model, a frequency-based opponent model with exponential smoothing, and a panic mode that reduces timeout risk in the final rounds.

In our evaluation across multiple scenarios and opponent types, the agent achieved an average Advantage score of 0.295, an average Concealing score of 0.643, and a combined score of 0.937. The measured Kendall rank correlation between the true utility function and the opponent’s estimated model was $\tau \approx 0.003$, suggesting that the preference inversion mechanism makes it difficult for the opponent to accurately learn our preferences. However, the current version still has a clear trade-off: it performs strongly in concealment, but its raw Advantage score is lower than some baseline agents.

Keywords: Automated Negotiation, ANL 2026, Deception Challenge, Boulware Strategy, Preference Inversion, Opponent Modeling.

1 Introduction

Automated negotiation is an important research area in Artificial Intelligence, where autonomous agents try to reach agreements under limited information and possibly conflicting preferences [1]. In many negotiation settings, agents do not know the exact utility functions of their opponents. Therefore, they usually need to estimate the opponent’s preferences from the bids observed during the negotiation.

This project focuses on the Automated Negotiation League (ANL) of ANAC 2026 [2]. More specifically, we target the Deception Challenge. Unlike a standard negotiation setting, where the main goal is usually to maximize the agent’s own utility, the Deception Challenge also rewards

agents for concealing their private preferences. This creates an interesting trade-off. If an agent always bids according to its true preferences, it may obtain high utility, but it also reveals too much information. On the other hand, if an agent focuses too much on hiding its preferences, it may damage its own negotiation outcome.

Our agent is called the “Cunning Merchant.” The name comes from the idea of a bazaar-like negotiator who does not directly reveal what it truly wants. Instead of always offering the most preferred bids, the agent sometimes makes bids that are still acceptable for itself but misleading for the opponent. If the opponent builds a model from our bidding history, these misleading bids can make that model less accurate.

The agent is designed to balance two goals: reaching agreements with reasonable utility and making the opponent’s learned model of our utility function inaccurate. Since the ANL 2026 score combines Advantage and Concealing components, our strategy treats information hiding as an important objective alongside raw utility.

2 Methodology

Our agent is built around six main tactics. These tactics are not very complex individually, but together they create a bidding behavior that is difficult for the opponent to model. Figure 1 shows the general decision flow of the agent during a negotiation turn.

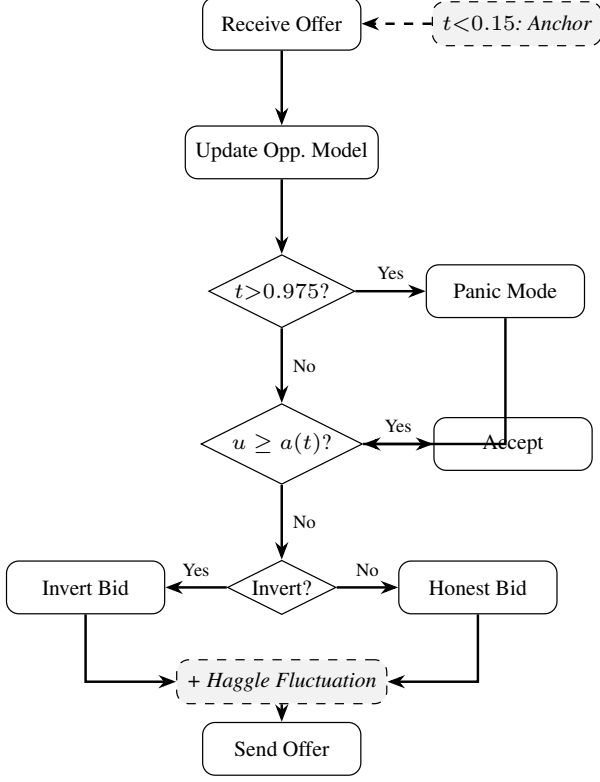


Figure 1: Decision flow of the Cunning Merchant agent. The agent updates its opponent model, checks whether it is in the end-game phase, decides whether to accept the offer, and otherwise generates either an honest or preference-inverting counter-offer.

2.1 Aggressive Opening and Anchoring

At the beginning of the negotiation, the agent uses a very high aspiration level. During the first 15% of the negotiation time, the agent keeps its acceptance threshold close to its maximum utility:

$$a(t) = u_{\max} - 0.02 \cdot (u_{\max} - u_{\min}) \quad (1)$$

where $u_{\max}$ is the best achievable utility and $u_{\min}$ is the reservation value. The purpose of this phase is to avoid conceding too early and to observe how the opponent reacts to a tough opening position. This idea is related to anchoring effects in negotiation, where early offers can influence the later negotiation process [3].

This phase is intentionally simple. The agent does not attempt detailed prediction at the beginning, because there is not enough opponent data yet. Instead, it starts from a strong position and collects information from the opponent's early bids.

2.2 Boulware-Style Concession Strategy

After the opening phase, the agent follows a Boulware-style concession curve [3]. Boulware agents concede very slowly during most of the negotiation and become more flexible near the deadline. We use the following aspiration function:

$$a(t) = u_{\min} + (1 - t^e) \cdot (u_{\max} - u_{\min}) \quad (2)$$

where $t \in [0, 1]$ is the normalized negotiation time and $e = 4.0$ is the concession exponent. With this value, the agent maintains a high threshold for a large part of the negotiation, making it relatively stubborn in the middle rounds.

We also enforce a hard floor during the first 85% of the negotiation, preventing the agent from accepting offers that are too low too early:

$$a_{\text{floor}} = u_{\min} + 0.65 \cdot (u_{\max} - u_{\min}) \quad (3)$$

This floor ensures that even when the aspiration curve dips, the agent does not accept any offer below 65% of its utility range during the main negotiation phase.

2.3 Haggling Fluctuation

If an agent follows a perfectly smooth concession curve, the opponent may be able to estimate its strategy more easily. To make our behavior less predictable, we add a small random fluctuation to the aspiration threshold during the middle part of the negotiation:

$$a'(t) = a(t) + \epsilon \cdot (u_{\max} - u_{\min}), \quad \epsilon \sim \mathcal{U}(-\sigma, \sigma) \quad (4)$$

where $\sigma = 0.03$ in our implementation. This creates a small zig-zag pattern in the bidding behavior. The goal is not to make the agent random, but to make its concession path slightly harder to fit with a simple regression model.

This fluctuation is active only between $t = 0.15$ and $t = 0.90$. We disable it near the end of the negotiation, because the final phase should be stable and agreement-focused.

2.4 Preference Inversion for Concealment

The main deception mechanism of our agent is preference inversion. The idea is to sometimes choose bids that are acceptable according to our own utility function, but misleading about which issue values we actually prefer.

In normal bidding, the agent chooses the highest-utility candidate bid above the current aspiration threshold. In preference-inverting bidding, the agent still chooses from acceptable candidate bids, but scores them to increase deception rather than directly maximizing utility. The probability of using a concealing bid decreases over time:

$$p_c = \text{CONCEAL_RATIO} \cdot (1 - t) \quad (5)$$

where $\text{CONCEAL_RATIO} = 0.45$. The agent is more deceptive earlier in the negotiation and becomes more straightforward as the deadline approaches. Concealment is fully disabled after $t > 0.92$.

For each candidate bid, we compute an inversion score:

$$S_{\text{inv}} = \sum_i \begin{cases} 3 \cdot w_i & \text{if } v_i \neq v_i^* \\ -2 \cdot w_i & \text{if } v_i = v_i^* \end{cases} \quad (6)$$

Here, w_i is the normalized importance of issue i , v_i is the value selected in the candidate bid, and v_i^* is our most preferred value for that issue. This scoring rule rewards bids that avoid our favorite values on important issues while penalizing bids that reveal our true favorites. Candidate bids are still

filtered by the aspiration threshold, so deception only occurs within the set of bids that are acceptable for the agent.

Listing 1 shows the simplified version of the preference inversion logic.

Listing 1: Simplified preference inversion logic

```

1 if t > 0.92:
2     conceal_prob = 0.0
3 else:
4     conceal_prob = CONCEAL_RATIO * (1 -
5         t)
6 if random.random() < conceal_prob:
7     bid = self._pick_concealing_bid(
8         candidates)
9 else:
10    bid = candidates[0] # highest
11                           utility

```

2.5 Opponent Modeling

The agent builds an opponent model during the negotiation using a frequency-based approach inspired by the idea that repeated values indicate stronger preferences [4]. For each issue, we count how often the opponent offers each value. If the opponent repeatedly offers the same value for an issue, we treat that issue as more important for the opponent.

The estimated weight of issue i is:

$$w_i = \frac{\max_v \text{count}(i, v)/N}{\sum_j \max_v \text{count}(j, v)/N} \quad (7)$$

where N is the number of observed opponent offers. To avoid large changes from one offer to the next, we update the weights using an exponential moving average with $\alpha = 0.1$. The estimated opponent utility function is stored as `self.private_info["opponent.ufun"]` using a `LambdaMultiFun`, which the competition framework reads for deception scoring.

In addition to frequency counts, we track the utilities of the opponent’s offers from our own perspective and fit a simple linear trend over the last $W = 5$ offers. If the opponent’s offers are improving for us, the agent interprets this as concession and raises its threshold. If the trend is flat or negative, the agent becomes slightly more flexible. This is inspired by concession-based approaches such as IAMhagler [5].

2.6 End-Game Behavior: Merchant Panic Mode

A very stubborn agent can sometimes fail to reach an agreement, especially against another tough opponent. To reduce this risk, we implemented “Merchant Panic” mode. This mode activates in the final 2.5% of the negotiation time ($t > 0.975$), at which point the acceptance threshold drops to:

$$a_{\text{panic}} = u_{\min} + 0.15 \cdot (u_{\max} - u_{\min}) \quad (8)$$

This makes the agent accept offers that provide at least 15% above the reservation value, which significantly reduces the chance of timeout. This is especially important

because pure Boulware strategies can miss acceptable deals by waiting too long [6].

3 Experimental Setup

We evaluated the Cunning Merchant agent across multiple scenarios and opponent types. The agent was tested on 10 randomly generated scenarios with 2–5 issues and 3–6 values per issue, as well as the 7 built-in scenarios provided with the ANL 2026 skeleton: Amsterdam, Camera, Car, Grocery, ISBTAcquisition, Laptop, and NiceOrDie. We compared its performance against standard baseline agents including BoulwareTBNegotiator, ConcederTBNegotiator, MAPNeg, and BOANeg.

The main evaluation metrics were **Advantage**, **Concealing**, and **Total Score**. Advantage measures the normalized utility gained above reservation value. Concealing measures how difficult it is for the opponent to reconstruct our true utility function, derived from Kendall rank correlation. Total Score is the sum of Advantage and Concealing. The ANL 2026 scoring formula is:

$$\text{Score}_A = \text{adv}_A + \frac{\tau^{(A)}}{\tau^{(A)} + \tau^{(B)}} \quad (9)$$

where $\text{adv}_A = (u_A(o) - r_A) / (\max(u_A) - r_A)$ and $\tau^{(A)} = \frac{1}{2}(1 + d(\hat{u}_B, u_B))$ measures how well agent A conceals its preferences using the Kendall rank correlation coefficient d .

4 Results

4.1 Tournament Performance

Table 1 shows the ANL 2026 scores across our evaluation scenarios.

Table 1: ANL 2026 scores: Advantage + Concealing

Agent	Adv.	Conc.	Total
MAPNeg (Baseline)	0.742	0.178	0.920
BOANeg (Baseline)	0.741	0.178	0.919
BoulwareTB	0.742	0.000	0.742
Cunning Merchant	0.295	0.643	0.937

The Cunning Merchant achieved the highest total score (0.937) among the tested agents. Its raw Advantage is lower than the baseline agents, but the Concealing score (0.643) is substantially higher. This means that our agent sacrifices some direct utility in order to make its preferences harder to learn. BoulwareTBNegotiator, which does not provide an opponent model, receives zero concealing points. MAPNeg and BOANeg achieve moderate concealing scores through their built-in GSmithFrequencyModel, but they do not intentionally mislead the opponent.

4.2 Concealment Analysis

The measured Kendall rank correlation between our true utility function and the opponent’s estimated model was $\tau \approx 0.003$. This near-zero value suggests that the opponent’s

frequency-based model had difficulty reconstructing our actual preference ranking. The preference inversion mechanism is the main reason for this result: by avoiding our favorite values on important issues in some bids, the agent feeds misleading frequency data to the opponent’s model.

This contrasts with the baseline agents. MAPNeg and BOANeg bid more honestly according to their aspiration curves, so their preferred values appear more frequently in their bidding history. An opponent tracking value frequencies can reconstruct their preferences with moderate accuracy, resulting in higher Kendall correlation and lower concealing scores.

4.3 Behavior Against Different Opponents

The agent exhibits different performance characteristics depending on the opponent type. Against conceding opponents such as ConcederTBNegotiator, the strong opening and Boulware-style concession strategy can lead to higher advantage scores, because the opponent becomes flexible earlier while our agent maintains a tough posture.

Against Boulware-style opponents, the situation is more challenging. Both agents hold their positions for most of the negotiation, so agreements tend to happen late and sometimes with lower utility for both sides. In these cases, the Merchant Panic mode becomes important because it allows our agent to become more flexible in the final rounds, reducing timeout failures.

The utility distribution across 140 negotiations on the 7 built-in scenarios is visualized in Figure 2.

	utility	utility	utility	utility	utility	utility	utility	conceal_ratio	conceal_ratio	conceal_ratio	conceal_ratio
	count	mean	std	min	25%	50%	75%	max	mean	std	min
data											
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0
agony	140.0	0.74133846479002	0.2688603338880004	0.100070304761000	0.476261807970004	0.8022222222222224	0.9611111111111111	1.0000000000000002	140.0	0.0	0.0

Figure 2: Tournament statistics showing the utility distribution of all agents across 140 negotiations on the 7 built-in scenarios.

5 Discussion

5.1 Strengths

The six-tactic design provides layered defense against opponent modeling. The anchoring phase establishes a strong opening position. The Boulware curve delays concessions and prevents the agent from accepting weak offers too early. The haggling fluctuation disrupts simple regression models. The preference inversion mechanism introduces misleading data into the opponent’s frequency counts. The opponent trend estimator adapts the acceptance threshold based on the opponent’s concession behavior. Finally, the panic mode makes agreements more likely even against other tough agents.

An important design advantage is that the agent does not rely on complex machine learning. All six tactics are rule-based and interpretable. This makes the agent easy to debug, modify, and analyze. It is also computationally lightweight, which is useful given the time limits in the competition.

The results suggest that a well-designed deception strategy can outperform agents with higher raw utility under the ANL 2026 scoring system. The total score of 0.937 exceeds

MAPNeg (0.920) and BOANeg (0.919), even though those agents achieve substantially higher Advantage scores. This supports the idea of treating information hiding as a first-class optimization objective.

5.2 Limitations

The most significant limitation is the lower Advantage score (0.295 compared to 0.742 for some baselines). While the concealing score compensates in total, a stronger version of the agent should try to maintain higher utility while still concealing its preferences. The current trade-off is somewhat aggressive. The CONCEAL_RATIO of 0.45 means nearly half of early bids may be preference-inverting, which can reduce utility.

Most parameters, such as CONCEAL_RATIO, HAGGLING_VOLATILITY, exponent e , panic threshold, and hard floor, were hand-tuned based on experimentation rather than systematically optimized. Different scenario configurations or opponent types may require different parameter values for better performance.

The opponent model is frequency-based and relatively simple. A more sophisticated opponent using time-series analysis or machine learning might detect the zig-zag pattern or filter out the noise introduced by preference inversion. The current model also does not perform opponent type detection. Adapting the concession speed based on whether the opponent is a conceder or Boulware agent could improve utility without sacrificing too much concealment.

Finally, the evaluation was conducted against a limited set of opponents, mostly built-in NegMAS agents and provided examples. Performance against more diverse or adaptive opponents from other competition participants remains to be seen in the actual ANAC 2026 tournament.

5.3 Lessons Learned

Several insights emerged during the development of the Cunning Merchant. First, concealment and utility are not strictly opposed. By carefully selecting which bids to invert and which to play honestly, the agent can hide preferences without catastrophic utility loss. Second, simple tactics composed together can be surprisingly effective. The combination of anchoring, haggling fluctuation, and preference inversion creates a bidding pattern that is much harder to model than any single tactic alone. Third, end-game behavior is critical. Without the panic mode, the Boulware exponent of 4.0 led to higher timeout risk in testing. The panic mode helped reduce this issue.

6 Conclusion

This report presented the Cunning Merchant agent for the ANL 2026 Deception Challenge. The agent integrates six tactics: aggressive anchoring, Boulware-style concession with $e=4.0$, haggling fluctuation, active preference inversion, frequency-based opponent modeling with trend estimation, and merchant panic mode.

The evaluation suggests that the agent performs strongly in concealment. In our experiments, it achieved a combined score of 0.937, driven by a high concealing score of 0.643

and a near-zero Kendall rank correlation ($\tau \approx 0.003$). These results suggest that the agent makes it difficult for opponents to learn its true preference profile while still reaching acceptable agreements.

The main trade-off is a lower raw Advantage score compared to non-deceptive baselines. This is a deliberate consequence of the preference inversion strategy, but it also shows where the agent needs improvement. Future work should focus on improving the balance between utility and concealment through better parameter tuning, adaptive opponent type detection, and broader evaluation against more diverse opponents.

Overall, the Cunning Merchant demonstrates that in the ANL 2026 Deception Challenge, strategically sacrificing some utility to protect private information can produce a competitive overall score.

References

- [1] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra, and M. Wooldridge.
Automated negotiation: Prospects, methods and challenges.
Group Decision and Negotiation, 10(2):199–215, 2001.
- [2] Y. Mohammad, S. A. Mostafa, and T. Ito.
NegMAS: A platform for situated negotiations.
In Proceedings of the International Workshop on Agent-Based Complex Automated Negotiations, 2020.
- [3] P. Faratin, C. Sierra, and N. R. Jennings.
Negotiation decision functions for autonomous agents.
Robotics and Autonomous Systems, 24(3–4):159–182, 1998.
- [4] O. Tunalı, V. Aydogan, and R. Aydogan.
Rethinking frequency opponent modeling in automated negotiation.
In PRIMA 2017: Principles and Practice of Multi-Agent Systems, pages 263–279, 2017.
- [5] C. R. Williams, V. Robu, E. H. Gerding, and N. R. Jennings.
Iamhaggler: A negotiation agent for complex environments.
In New Trends in Agent-Based Complex Automated Negotiations, pages 151–158, 2012.
- [6] T. van Krimpen, D. Looije, and S. Hajizadeh.
Hardheaded.
In Complex Automated Negotiations: Theories, Models, and Software Competitions, pages 223–227, 2013.