

Hallucinators: A Preference-Concealing Negotiation Agent for the ANL 2026 League

Diya Panikkassery Sudheer Kumar

Department of Computer Science, Leibniz University Hannover

diya.panikkassery.sudheer.kumar@stud.uni-hannover.de

Gayatri Rajeev

Department of Computer Science, Leibniz University Hannover

gayatri.rajeev@stud.uni-hannover.de

Kevin Antony

Department of Computer Science, Leibniz University Hannover

kevin.antony@stud.uni-hannover.de

Abstract

We present *Hallucinators*, our agent for the ANAC 2026 Automated Negotiation League (ANL). The ANL 2026 score rewards two equally weighted objectives: securing a high-utility agreement (*Advantage*) and hiding one’s own preferences from the opponent (*Concealing*). *Hallucinators* addresses both through six complementary mechanisms: (i) a three-phase Boulware concession schedule with *opponent-adaptive exponents* that stay Boulware longer against hardliners, (ii) a distribution-based frequency opponent model with dynamic issue-weight refinement and behavioral profiling, (iii) an active preference-masking strategy with deterministic, reproducible offer selection, (iv) extreme conflict detection with hardliner-advantage guards to manage adversarial negotiation domains, (v) a dynamic weight-modulation layer that feeds opponent behavioral profile scores directly into the multi-objective scoring weights, and (vi) a negative-correlation cold-start prior that improves early-game candidate ranking before opponent data is available. In a local Cartesian tournament against four standard baselines across all seven provided scenarios, *Hallucinators* achieved a mean final score of **1.405** (Advantage: 0.786, Concealing: 0.620), outperforming all competitors on both sub-components.

1 Introduction

The ANL 2026 competition evaluates bilateral negotiation agents on a *combined* score:

$$\text{Score} = \underbrace{(u(\omega) - r)}_{\text{Advantage}} + \underbrace{C}_{\text{Concealing}},$$

where $u(\omega)$ is the agent’s utility for the agreed outcome ω , r is its reservation value, and C is the fraction of the Kendall- τ ranking point not captured by the opponent’s frequency model of the agent. Maximising both objectives simultaneously is challenging: high-utility offers tend to be concentrated near the agent’s ideal point, which is precisely where the opponent’s model learns to look.

Our approach, *Hallucinators*, manages this tension by separating offer generation into a *utility-safe candidate set* and a *masking-optimal selection* within that set. Concession follows a three-phase Boulware schedule. A robust acceptance strategy with extreme-conflict rescue prevents sub-optimal agreements early on and deadlocks later, while a two-level candidate filter and RUFL-inspired phase policy select offers that balance own utility with opponent acceptance.

2 The Design of Hallucinators

Hallucinators extends the `SAOCallNegotiator` base class from `NegMAS`, giving full control over acceptance and offering through a single `__call__` entry point. An earlier version extended `BOANegotiator`; the switch eliminates the unused `TimeBasedOfferingPolicy`, `ACNext`, and `GSmithFrequencyModel` components that were initialised but bypassed by our override, reducing startup overhead in large outcome spaces. All tunable constants are collected in a frozen `NegotiationConfig` dataclass (approx. 60 parameters), making the agent easily reproducible and tunable.

Offer management. At the start of each negotiation, `on_preferences_changed` enumerates all outcomes above the reservation value, sorts them by own utility, and caches them in `_outcomes`. The agent also precomputes a `_frontier_cache` of three named utility bands (`top_utility`, `safe_compromise`, `masked_rotation`) used by the late-game offer selection routines, avoiding repeated scans of the full outcome list during the negotiation. Additionally, all own utilities are stored in a sorted list to support fast percentile queries. The estimated opponent utility function is published in `private_info` (required by the ANL scoring infrastructure).

Three-phase concession. The target utility $T(t)$ at relative time $t \in [0, 1]$ follows:

$$\begin{aligned} t < 0.55 &\Rightarrow T = 0.92 u_{\max} \quad (\text{early, aggressive}), \\ 0.55 \leq t < 0.85 &\Rightarrow T = 0.92 u_{\max} - 0.17 u_{\max} \cdot \left(\frac{t-0.55}{0.30} \right)^{e_{\text{mid}}} \quad (\text{mid, Boulware concession}), \\ t \geq 0.85 &\Rightarrow T = 0.75 u_{\max} - 0.11 u_{\max} \cdot \left(\frac{t-0.85}{0.15} \right)^{e_{\text{late}}} \quad (\text{late, final concession}), \end{aligned}$$

with a floor of $\max(r, 0.64 u_{\max})$. The exponents ($e_{\text{mid}}, e_{\text{late}}$) are no longer hardcoded but computed by `_concession_exponents()` at each step based on the opponent’s *hardliner score* and *concession rate*: by default (2.4, 1.4); against detected hardliners (3.2, 2.0); against detected conceders (2.0, 1.8). Staying Boulware longer against hardliners prevents unnecessary concessions when the opponent is unlikely to reciprocate. An opponent-responsive adjustment additionally nudges T up or down based on observed concession rate and model confidence once 45% of time has elapsed (confidence threshold raised to 0.45 to require stronger evidence). On the final step, T is relieved by 2.5% $u_{\max}$ to secure last-second agreements.

Dynamic weight modulation. The multi-objective scoring weights ($w_{\text{own}}, w_{\text{acc}}, w_{\text{par}}, w_{\text{mask}}$) are computed at each step by `_weights(relative_time)`. Starting from the phase-specific base vectors (Section 5), the method applies opponent-profile adjustments: a high hardliner score boosts w_{acc} (we prioritise acceptable bids); a high conceder score boosts w_{par} (we exploit Pareto gains); a high threshold-acceptor score modestly lifts w_{own} . All adjustments are gated by a minimum confidence threshold to prevent noisy early-game profiles from distorting the weights.

3 Concealing Bidding Strategy

Our masking pipeline operates in three layers.

Issue-cost estimation. At initialisation we estimate each issue’s *cost* — how much own utility varies across its values. Issues with low cost are “cheap” (we can vary them freely); issues with high cost are “expensive” (fixing them near our optimum is essential).

Cold-start opponent prior. Before any opponent offers are received, the agent cannot rely on frequency counts to estimate opponent utility. Rather than returning a flat 0.5 for all outcomes (which makes every early candidate score identically), Hallucinators uses a mild *negative-correlation prior*:

$$\hat{u}_{\text{opp}}^{\text{prior}}(\omega) = 0.7 - 0.3 s_{\text{own}}(\omega),$$

where s_{own} is the agent’s own normalised utility. This reflects the typical opposing-preference structure in zero-sum-like domains and meaningfully differentiates candidates even in the first round.

Deterministic offer selection. The masked-window selection now always returns the *top-scored* offer from the deterministically sorted pool, rather than using the earlier $(\text{step} \times 5 + \lfloor t \times 100 \rfloor) \bmod |\text{window}|$ index pattern. This eliminates run-to-run score variance caused by negotiation-step indexing and makes tournament results fully reproducible.

Decoy signalling. During the early and mid phases ($t < 0.78$) each candidate offer is scored by a *decoy signal*:

$$s_{\text{decoy}} = 0.65 \cdot \text{cheap_stability} + 0.35 \cdot \text{expensive_variety},$$

where *cheap_stability* rewards keeping cheap-issue values the same as the previous offer (signalling *fake* resistance), and *expensive_variety* rewards changing expensive-issue values (hiding the high-cost dimension).

Masked window selection. After scoring all candidates, we form a *masked window*: the subset of top candidates whose own utility falls within a narrow slack band (3.5% early, 7% late) around the best candidate. Within this window, offers are re-ranked by masking score — issue novelty, exact-repeat avoidance, and recent-value reuse — to maximally confuse the opponent’s frequency model.

Two-level candidate filtering and RUFL policies. Before multi-objective scoring, the candidate pool undergoes a two-level selection pass. The filter activates from mid-game onwards ($t \geq 0.45$, lowered from 0.60; the opponent model is sufficiently reliable after ≈ 8 offers, which typically arrive before 45% of time elapses). First, we preserve a utility-safe band (within 3.5% of the best available utility), then we rank these safe candidates by their estimated acceptance likelihood, keeping the top 60%. Additionally, RUFL-inspired phase-specific policies are applied late in the negotiation. Near the absolute deadline ($t \geq 0.94$), the regular selection is overridden entirely to explicitly propose the best offer previously made by the opponent, or bids with extreme opponent-acceptance likelihoods, bypassing masking completely to avert timeouts.

4 Acceptance Strategy

Acceptance is checked in the following priority order.

1. **Reservation guard.** Reject any offer below r .
2. **Extreme conflict rescue.** If the domain is tiny and the negotiation is stalled, we attempt an extreme-conflict rescue. We accept the opponent’s best offer seen so far if it is within 2% of our own rescue compromise. However, if the opponent is identified as a strong hardliner (requiring at least 8 prior offers for reliable classification), this rescue is bypassed to prevent exploitation.
3. **ACcombi** (Baarslag et al., 2010). At $t \geq 0.99$, accept if the offer utility is at least the window-average of recent opponent offers. A minimum window of `MIN_OFFER_HISTORY` = 3 entries is enforced before this rule fires, preventing acceptance based on a single noisy offer.
4. **Target utility.** Accept if $u(\text{offer}) \geq T(t)$.
5. **Last-chance.** On the final step ($t \geq 0.97$, aligned with `LATE_GAME_THRESHOLD`), accept the current offer if it is at least as good as the best offer the opponent has made so far and meets $0.98 T(t)$.
6. **Counter-offer margin.** Accept if $u(\text{offer}) \geq (0.995 - 0.08 t) \cdot u(\text{planned})$.
7. **Late-game floor.** At $t > 0.97$ accept any offer above the emergency floor $0.55 u_{\max}$ (raised to $0.62 u_{\max}$ if the hardliner advantage guard is active).

5 Opponent Model

Distribution-based frequency model. The `DistributionAnalyzer` maintains per-issue weighted frequency counts $C_i(v)$ of opponent offers. The estimated utility of an outcome ω is a weighted sum of per-issue value

scores:

$$\hat{u}_{\text{opp}}(\omega) = \sum_i w_i \cdot \left(\frac{C_i(\omega_i) + \lambda}{C_i^* + \lambda} \right)^{0.35},$$

where $\lambda=1$ is a Laplacian smoothing factor, C_i^* is the maximum count over values of issue i , and w_i is the estimated issue weight. The power exponent 0.35 compresses high-frequency signals to reduce overconfidence. Earlier offers have higher model weight; repeated identical offers are down-weighted exponentially (damping factor 0.48, minimum weight 0.18).

Issue-weight update. Every $k=4$ offers we compare the L_1 distance between the previous and current window distributions. Issues that remain *stable* in a window where the opponent conceded on at least one other issue are inferred to be high-priority; their weights are boosted by $0.08(1 - t)$ and the vector is renormalised.

Behavioral profiling, weight adaptation, and acceptance likelihood. Beyond frequency modeling, the agent maintains an active four-component behavioral profile of the opponent, estimating whether they act as a *hardliner*, a *conceder*, a *fixed-threshold acceptor*, or a *mirror strategist*. These scores directly feed into `_weights()` to shift the multi-objective scoring weights at each step: a hardliner opponent raises w_{acc} (we must propose bids the opponent will actually accept); a conceder opponent boosts w_{par} (we can afford to explore Pareto-optimal bids); a threshold-acceptor opponent modestly increases w_{own} . All adjustments require a minimum model confidence to avoid distortion from early noisy data. The profiling scores also modulate concession aggressiveness and extreme-conflict rescue guards. Furthermore, they directly inform our estimated acceptance likelihood (s_{acc}) for candidate offers. The acceptance likelihood integrates the frequency-based opponent score with the offer’s structural similarity to past opponent bids, while applying dynamic bonuses for hardliners (who demand high similarity) and patience-based adjustments as the deadline approaches.

Opponent deadline forecasting and Lookahead. To evaluate candidate offers under uncertainty, we predict the best utility the opponent may offer by the deadline, adapting a utility-fit approach from Team 271. We compute a median over pairwise concession slopes to establish a robust trend line, and scale it by the remaining time. This forecast drives a one-step lookahead tiebreak score ($s_{\text{lookahead}}$) and determines our confidence in the opponent model.

Multi-objective candidate scoring. Each candidate offer is scored as:

$$S = w_{\text{own}} s_{\text{own}} + w_{\text{acc}} s_{\text{acc}} + w_{\text{par}} s_{\text{par}} + w_{\text{mask}} s_{\text{mask}} + 0.04 s_{\text{deadline}} + 0.035 s_{\text{decoy}} + 0.025 s_{\text{lookahead}},$$

with phase-specific weight vectors ($w_{\text{own}}, w_{\text{acc}}, w_{\text{par}}, w_{\text{mask}}$): early (0.74, 0.06, 0.04, 0.16), mid (0.60, 0.16, 0.10, 0.14), late (0.48, 0.28, 0.16, 0.08).

6 Evaluation

We evaluated Hallucinators in a Cartesian round-robin tournament against four standard NegMAS baselines across all seven provided scenarios (Amsterdam, Camera, Car, Grocery, ISBTAcquisition, Laptop, NiceOrDie), yielding 140 negotiations per agent.

Table 1: Mean tournament scores (Advantage + Concealing = Score) from a Cartesian round-robin across all seven provided scenarios (140 negotiations per agent). Higher is better.

Agent	Advantage	Concealing	Score
Hallucinators (ours)	0.786	0.620	1.405
BOANeg	0.671	0.616	1.287
MAPNeg	0.671	0.616	1.287
SimpleNegotiator	0.477	0.548	1.025
BoulwareTBNegotiator	0.671	0.000	0.671

Advantage. Hallucinators achieves a mean Advantage of **0.786**, compared to 0.671 for both BOANeg and MAPNeg. Several improvements contribute to this gain. Adaptive concession exponents (FIX 10) allow the agent to stay Boulware longer against hardliners, avoiding premature concessions. The two-level candidate filter, now activated from $t \geq 0.45$ rather than 0.60 (FIX 6), makes opponent-model-guided candidate selection effective significantly earlier in the negotiation. The hardliner advantage guard now requires at least 8 offers before firing (FIX 8), producing more reliable conflict-rescue decisions. Dynamic weight modulation (FIX 4) raises w_{acc} against hardliners and boosts w_{par} against conceders, focusing bid selection on what the specific opponent is actually likely to accept.

Concealing. The Concealing score of **0.620** reflects the effectiveness of the decoy-signal, masked-window, and deterministic selection strategies in confusing the opponent’s frequency model. Deterministic offer selection (FIX 1) removes the run-to-run variance introduced by the previous modulo index pattern, producing a more consistent concealing pattern across negotiations. The negative-correlation cold-start prior (FIX 3) also contributes: early offers are now selected against a meaningfully differentiated opponent estimate rather than a flat 0.5, helping choose bids that both extract utility and obscure preferences from the first round. BoulwareTBNegotiator scores 0.000 because it does not publish an `opponent_ufun` in `private_info`, and the tournament infrastructure therefore cannot score its opponent model.

7 Lessons and Suggestions

1. **Deterministic selection is essential for reproducibility.** The original masked-window code indexed the offer pool with a step-based modulo, introducing run-to-run score variance. Switching to a deterministically sorted pool (always choosing index 0 of the highest-scoring candidates) eliminated this variance and produced more consistent tournament results.
2. **A cold-start prior meaningfully differentiates early candidates.** Returning a flat 0.5 for all outcomes before any opponent data is available caused every early candidate to score identically, making the first few bids effectively random. Replacing this with a mild negative-correlation prior ($0.7 - 0.3 s_{\text{own}}$) gave the scoring function meaningful signal from round one.
3. **Adaptive concession exponents improve Boulware robustness.** Hardcoded exponents (2.4 mid, 1.4 late) treated all opponents identically. Computing exponents from the observed hardliner score and concession rate — staying Boulware longer against hardliners — prevented unnecessary concessions against opponents who do not reciprocate.
4. **Opponent profiles must feed into behaviour, not just be computed.** An early version of the agent computed hardliner, conceiver, mirror, and threshold-acceptor scores but never acted on them. Wiring these scores into `_weights()` to dynamically shift component weights meaningfully improved both Advantage and Concealing.
5. **Guarding against hardliners requires sufficient evidence.** The hardliner advantage guard was a crucial contributor to improved Advantage. Raising the minimum-evidence requirement to 8 offers (from 4) prevented the guard from firing on noisy early data, making the conflict-rescue suppression both safer and more effective.
6. **Two-level candidate selection balances utility and acceptance.** Filtering the top candidates by a strict own-utility slack (3.5%) and then keeping the top 60% ranked by opponent acceptance likelihood ensures that the agent only offers highly acceptable bids without sacrificing Advantage. Activating this filter from $t \geq 0.45$ rather than 0.60 exploits the opponent model earlier and improves mid-game bid quality.
7. **Require strong evidence before acting on opponent models.** The responsive concession and the adaptive weight modulation both apply confidence thresholds. Raising `RESPONSIVE_CONFIDENCE_THRESHOLD` from 0.30 to 0.45 reduced erratic mid-game behaviour caused by premature activation on noisy early data.

Conclusions

Hallucinators demonstrates that the ANL 2026 dual objective of maximising agreement utility while concealing preferences can be addressed through a unified pipeline. By separating candidate generation (utility-safe) from candidate selection (masking-optimal), coupling behavioral profiling directly into both the scoring weights and the concession schedule, and enforcing deterministic, reproducible offer selection, the agent achieves a mean score of **1.405** (Advantage: 0.786, Concealing: 0.620) in a local Cartesian tournament — a **+2.4%** gain over the previous version. Key engineering improvements include opponent-adaptive concession exponents, an earlier-activated two-level candidate filter, a negative-correlation cold-start prior, and stricter evidence requirements for the hardliner advantage guard. Future work will explore learned models of opponent types and multi-round evidence aggregation.

References

- [1] T. Baarslag, K. Hindriks, C. Jonker, S. Kraus, and R. Lin. The first automated negotiating agents competition (ANAC 2010). In *New Trends in Agent-Based Complex Automated Negotiations*, 2012.
- [2] O. Tunali, T. Aydogan, and M. Sanchez-Anguix. Rethinking frequency opponent modeling in automated negotiation. In *Proc. PRIMA 2017, LNCS*, 2017.