

AnlOmegaNegotiator

A Risk-Safe, Concealment-Aware Hybrid Negotiator for ANL 2026

Genchan Omega
ANL 2026 Submission

Abstract

We present AnlOmegaNegotiator, a bilateral negotiation agent for the ANAC 2026 Automated Negotiation League. The agent combines an adaptive frequency model, hypothesis-based opponent reasoning, Pareto/Nash guidance, concealment-aware bid diversification, style-sensitive concession, and strict individual rationality. The final revision was driven by the frozen online tournament and controlled local ablation. It removes a negative-Advantage acceptance failure and separates the model used for strategic bidding from a Smith frequency model reported after negotiation for scoring. Across all seven supplied scenarios, both utility profiles, and six official-style opponents, the final change raises the score proxy from 1.0242 to 1.0338, raises the model component from 0.5093 to 0.5189, and improves the minimum from 0.4821 to 0.5131 without changing mean Advantage or agreement rate.

1 Problem and Design Objectives

ANL 2026 rewards both agreement quality and preference uncertainty. A useful agent must obtain utility above its reservation value, infer the opponent’s ranking from a short offer sequence, and avoid making its own ranking obvious. These objectives conflict: repeated high-utility offers reveal preferred issue values, aggressive accommodation loses Advantage, and excessive concealment causes deadlock.

Our design follows four constraints. First, no agreement may be worse than disagreement. Second, aspiration controls the feasible utility band before opponent estimates influence ranking. Third, opponent modeling must remain lightweight enough for 100–2000 negotiation rounds. Fourth, every additional mechanism must survive broad ablation rather than only improve one scenario.

2 Agent Architecture

The submission package exposes AnlOmegaNegotiator from module `submission.anl_agent.entry`. It uses the tuned HybridNegotiator policy and has six main stages.

2.1 Outcome preprocessing and aspiration

At initialization, the agent enumerates or samples up to 40,000 outcomes, evaluates them with its own utility function, removes outcomes at or below the reservation value, and stores the remaining outcomes in descending utility order. A piecewise aspiration curve defines a target utility at each relative time. Candidate bids are drawn from bands around that target, so opponent accommodation cannot silently force an irrational concession.

2.2 Adaptive opponent model

For every issue value, the agent tracks weighted frequency, recency, and value transitions. Issue importance combines value concentration with stickiness: issues the opponent changes rarely receive more weight. A small hypothesis set adds mirror and issue-focused priors during sparse early observations. The resulting additive model is used internally for bid ranking and frontier guidance.

2.3 Offer selection

Each candidate receives a composite score based on self-utility, estimated opponent utility, reciprocity, recent-offer alignment, diversity, balance, Pareto-frontier membership, Nash proximity, and deadlock signals. Late in the negotiation, joint utility receives extra weight. Candidate pools also include outcomes near recent and best received offers when reciprocal or stubborn behavior is detected. This gives the agent multiple routes to agreement without abandoning its aspiration floor.

2.4 Concealment

The agent avoids repeatedly exposing its highest-valued issue settings. Diversity rewards new combinations and distance from recent bids; reveal penalties discourage repeated high-ranked values; an early camouflage score prefers less diagnostic mid-valued settings. These mechanisms decay over time so privacy pressure does not block a late agreement.

2.5 Acceptance and individual rationality

Acceptance compares the current offer with aspiration, the next planned bid, and a projected future offer. Plateau, reciprocity, and hardline signals can relax the threshold near the deadline. Every acceptance path first enforces

$$u_{self}(o) \geq r_{self},$$

and the threshold is clamped again after all adaptive discounts. This invariant prevents the negative-Advantage tail found in the online result. An offer equal to reservation remains admissible because it cannot be worse than disagreement and may preserve the model component of the ANL score.

2.6 Tiny-domain rescue

One-issue domains such as NiceOrDie receive dedicated handling. The agent detects repeated incompatible extremes, estimates the central compromise using both utilities, and offers or accepts it near the deadline. Generic multi-issue search otherwise oscillates and frequently misses this agreement.

3 Final Model Separation

The previous implementation used one estimated utility function both for strategic bidding and as the final reported opponent model. Changing its parameters could improve ranking accuracy while simultaneously changing bids and harming TFT or hardline performance. The final implementation separates these responsibilities.

During negotiation, the established adaptive model remains unchanged and is used for Pareto and Nash calculations. At `on_negotiation_end`, the reported model is replaced by a Smith-style frequency model:

$$\hat{u}_{Smith}(o) = \frac{1}{m} \sum_{i=1}^m \frac{c_i(o_i)}{\max_v c_i(v)},$$

where m is the number of issues and $c_i(v)$ is the observed frequency of value v for issue i . The replacement occurs only after bargaining ends, so it cannot change proposals, acceptance, or agreement. It improves the ranking estimate used by the ANL scoring procedure while preserving the tested hybrid strategy.

4 Online Diagnosis

The frozen tournament 19190 contained 40 agents, eight configurations, and 25,600 negotiations. CodexAgentAnl ranked 35th with mean score 4,000.57 over 1,280 negotiations. Table 1 shows the relevant distribution.

Table 1: Frozen online result used for final diagnosis.

Rank	Score	Min	Q1	Median	Q3	Max	Time (ms)	Exceptions
35/40	4000.57	-3971	2488	2944	5890	10000	542.32	0

The negative minimum reproduced locally when reservation values were nonzero: adaptive discounts were applied after an initial reservation clamp. A 48-negotiation stress test found 25 below-reservation agreements with minimum Advantage -0.4701 . After the invariant fix, violations fell to zero. The low median also motivated profile-symmetric evaluation rather than testing only the first utility function of each scenario.

5 Experimental Method

Development used repeated one-change-at-a-time evaluation. The final round tested thirteen conditions: three Smith blend levels, higher self and opponent weights, lower balance and Nash weights, stricter acceptance, wider near-best acceptance, lower late aspiration, lighter reveal and camouflage pressure, and a smaller recent-offer pool.

The search suite included Camera, Grocery, ISBTAcquisition, and Laptop; both profiles; four official-style opponents; generated time-based, TFT, Tough, and MiCRO opponents; and nonzero-reservation stress domains. Every candidate was rejected if it violated reservation or degraded the lower tail materially. The accepted search configuration was then ablated on all seven scenarios, both profiles, and six official-style opponents. Strategy-weight changes that looked useful in the reduced search were removed after the full evaluation.

Table 2: Final fixed-scenario validation.

Metric	Previous	Final	Change
Official score proxy	1.024221	1.033820	+0.009599
Advantage	0.514926	0.514926	0.000000
Model component	0.509295	0.518894	+0.009599
Minimum score	0.482065	0.513131	+0.031066
10th percentile	0.665476	0.665476	0.000000
Agreement rate	100%	100%	0 pp
Reservation violations	0	0	0

The older eight-opponent official benchmark independently improved from 1.068104 to 1.076490. The final revision therefore improves model quality and the minimum score without trading away mean Advantage, agreement, or safety.

6 Ablation Findings

Several plausible alternatives were rejected. Replacing the hybrid policy with the template BOA/MAP architecture improved model quality on a narrow symmetric suite but reduced the broader objective from 0.8173 to 0.7579 and obtained zero utility against Tough in that stress set. Increasing self weight reduced both the official proxy and generated utility. Lowering Nash weight, aspiration, or reveal pressure also failed controlled comparisons. Lower balance weight and higher opponent weight looked useful in the reduced round-15 suite, but the full seven-scenario ablation found weaker stress behavior and no reliable gain over Smith reporting alone. These negative results motivated the conservative final configuration.

7 Complexity and Deployment

Outcome preprocessing costs $O(\min(|\Omega|, 40000))$. Per round, model updates are linear in the number of issue values and bid ranking is capped at 300 candidates. Pareto/Nash guidance uses a cached reference set and refreshes every two model updates. Value-frequency and reuse maps are maintained incrementally, avoiding full-history rescans for each candidate. The final ZIP contains only the submission package and uses NegMAS as its sole external dependency.

8 Limitations and Reproducibility

Local scores are proxies because final opponents and the eighth online configuration are not public. The supplied scenarios nevertheless cover tiny, small, and large outcome spaces, and both utility profiles were tested to avoid role bias. Results are recorded in `results/final_tuning_round15_results.json` and `results/final_validation_results.json`. The tested class is `AnlOmegaNegotiator` in module `submission.anl_agent.entry`; ZIP and directory imports were checked independently under Python 3.14 and NegMAS 0.15.7.

9 Conclusion

AnlOmegaNegotiator is a conservative final system: it combines adaptive opponent learning, concealment-aware utility bands, frontier guidance, style-sensitive endgame behavior, and strict individual rationality. The last revision does not replace a stable bidding strategy with a locally optimized one. Instead, it fixes the observed negative tail and improves the reported opponent ranking after negotiation, producing a measurable score increase with minimal behavioral risk.