

Anchor: A Score-First Negotiation Agent for ANL 2026

David Izhaki Itai Kahana
davidizhakinew@gmail.com Itaygil135@gmail.com
Bar-Ilan University

June 2026

Abstract

We describe *Anchor*, our entry to the ANAC 2026 Automated Negotiation League (ANL) [4]. The competition scores an agent both on the deal it reaches and on how well it hides its own preferences while modelling its opponent’s. Because the tournament ranks by absolute score rather than by margin over any single opponent, Anchor maximises Advantage: it bids firmly, closes deals instead of holding out, and spends its effort on modelling the opponent accurately, deceiving only when that costs nothing. Across about 30 NegMAS [1] opponents on the provided and randomly generated domains it scores positively against every type, averaging 1.52. We also implemented three deliberate concealment mechanisms and measured each to be counter-productive: under this scoring rule, the way to hide your preferences is to bid firmly and model the opponent well, not to add noise to your offers.

1 Approach

Two features of the ANL 2026 score, Advantage plus Concealing, shape the design. Advantage is a full additive term, while the contestable part of the shared Concealing point is small. The tournament also ranks agents by their absolute mean score, not by their margin over a given opponent: a no-deal yields only the Concealing term (about 0.5), whereas a balanced deal yields about 1.0, so closing a deal is worth far more than holding out to “win” a single negotiation. The free half of Concealing, modelling the opponent well, also sharpens our own offers; its costly half, making the opponent model us badly, rarely pays for itself (Section 6). Anchor is therefore a Bidding–Opponent–modelling–Acceptance (BOA) agent that bids firmly, banks Advantage by closing deals, and invests in modelling rather than deception.

2 The Design of Anchor

Anchor is a single `SAOCallNegotiator` implementing the three classical BOA components [2] (bidding, opponent modelling, and acceptance), described below. All state is rebuilt at the start of each negotiation, respecting the no-cross-negotiation-memory rule; the agent writes no files and prints nothing.

A guiding invariant is that we emit an opponent-utility estimate on *every* turn (Section 3): the scorer reads it to compute τ_A , and a missing or degenerate estimate forfeits the entire Concealing point. This matters in practice: our evaluation in Section 6 shows that fixing a corner case in which this estimate could become degenerate raised our mean score by ~ 0.1 .

3 Opponent Model

We use a frequency-based additive model, reconstructed every round as a `LinearAdditiveUtilityFunction` (a per-issue value table times an issue-weight vector). It recovers the *ranking* of outcomes, which

Table 1: Opponent-modelling accuracy (Kendall- τ -based, full outcome space), mean over the model-capable opponents. “Recency” was our initial scheme; “Early” is the shipped model. τ_A : how well we model them (higher is better); τ_B : how well they model us. Share = our Concealing fraction.

Time weighting	τ_A (we model them)	τ_B (they model us)	Concealing share
Recency (initial)	0.479	0.55	0.484
Uniform	0.536	0.55	0.496
Early (shipped)	0.571	0.56	0.502

is what the Kendall- τ scoring rewards, rather than cardinal utilities or the opponent’s reservation value.

Value scores. Within an issue, a value the opponent offers more often is assumed better for them (normalised frequency).

Issue weights. Issues whose value the opponent rarely changes are assumed more important (stability), with a uniform prior until enough offers are seen.

Early-time weighting (our main modelling result). The most effective change was deciding *when* to trust offers. An opponent’s *early* offers sit near its ideal and reveal its true preferences. Its *late* offers are concessions: they show what it will accept, not what it wants. We therefore weight earlier offers more heavily, the opposite of the recency weighting we used at first. The effect on modelling accuracy is large (Table 1): mean τ_A rose from 0.48 to 0.57, and against the hardest-to-model baseline (an opponent that assumes our preferences are the inverse of its own) our τ_A rose from 0.27 to 0.59. This moved us from *losing* the Concealing split on average (0.48) to *parity* (0.50).

Non-degeneracy guard. An empty model assigns equal utility to all outcomes, for which Kendall- τ is undefined; the scorer then assigns us zero accuracy and we forfeit our entire Concealing share. This bites in an otherwise normal case: when we open and the opponent accepts immediately, we are scored before having folded in any opponent offer. A negligible deterministic gradient guarantees that the emitted model is never perfectly flat; this alone raised our mean tournament score by ~ 0.1 , recovering the Concealing point against the many opponents that accept early.

We also tested a Bayesian variant (a posterior over rank-weight hypotheses) and a switch to NegMAS’ weight-learning frequency models; on τ_A , both were within noise of the shipped model. Consistent with the literature [3], Kendall- τ needs only the ranking, which our frequency model already recovers from ~ 100 offers, so we retained it.

4 Concealing Bidding Strategy

Our bidding decouples *how much* to concede from *which* bid to make.

How much: a floored Boulware target. We set a time-based aspiration that starts at our maximum utility and concedes slowly (a firm Boulware exponent). The target does not fall all the way to the reservation value: it is floored at a fair fraction of our utility range. This is the **anti-exploitation floor**. A naive Boulware that concedes down to its reservation value is catastrophic against a firm opponent: in early testing against a strong, hand-built baseline

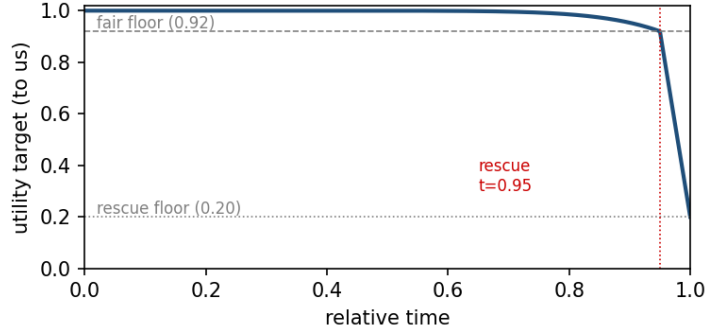


Figure 1: Our utility target over time. We hold firmly above the anti-exploitation fair floor (0.92 of our range), then engage a short end-game rescue (after $t = 0.95$) that concedes toward a rescue floor (0.20) to close a deal. Conceders agree before the rescue engages, so we still extract from them; only a firm opponent pushes us into the rescue.

opponent, a naive conceding agent was walked all the way down to the opponent’s ideal, *losing* by 0.48 on average. Flooring the concession (Figure 1) removed the exploitation entirely.

Which bid: opponent-aware selection. Among all outcomes worth at least the current target to us, we offer the one our opponent model rates highest (a Nash-style product of our and their estimated utility). This raises agreement probability at no cost to our own utility, because every candidate already clears our floor.

Closing deals: the score-first end-game. Here our objective (absolute score, not margin) reshapes the design. Holding firm to the end maximises our *margin* over a firm opponent, but it produces a no-deal, which scores only the Concealing term. Since a balanced deal scores roughly twice a no-deal, we deliberately concede within a short end-game “rescue” window toward (but not all the way to) the reservation value, and we *capture* the best offer the opponent has shown us as soon as it clears a high bar. Re-tuning the rescue floor for score rather than margin raised our mean score against firm opponents from 1.07 to 1.19, at no cost against conceders (who concede before the rescue engages). This trades the head-to-head “win” over a firm opponent (margin $+0.08 \rightarrow 0.00$) for a higher *absolute* score everywhere, which is what the tournament rewards.

5 Acceptance Strategy

Acceptance composes with bidding. We accept the opponent’s offer when, in order: (i) it is at least our reservation value (a hard floor: we never accept worse than walking away); (ii) it is at least as good for us as the bid we would make next (the standard **ACNext** rule); (iii) it is the best the opponent has yet offered and clears a high “capture” threshold, so we do not hold out for our maximum and lose a near-best offer that may not return; or (iv) in the end-game, it is the opponent’s best *fair* offer, which we bank rather than risk a Concealing-losing no-deal. Because the planned next bid respects the fair floor before the end-game, rule (ii) accepts only genuinely good offers early; the relaxation that closes deals is confined to the end-game.

6 Evaluation

Method. We built a local evaluation harness that runs Anchor against a roster of ~ 30 real NegMAS negotiators (conceders, hard-liners, tit-for-tat, time-based, and strong hybrids), across the seven provided domains and a batch of randomly generated domains spanning cooperative

Table 2: Our mean absolute score (Advantage + Concealing) by opponent group, over ~ 30 opponents $\times$ local + generated domains $\times$ both orders. A no-deal scores ≈ 0.5 ; a balanced deal ≈ 1.0 .

Opponent group	Our mean score	Deal rate
Conceders (Conceder, Linear, Boulware, Nice, CAB/WAB, ...)	1.6 – 1.84	high
Strong non-modellers (Tough, MiCRO, Hybrid, TopFraction)	~ 1.0	low [†]
Frequency modellers (BOANeg, MAPNeg)	1.10	high
Inverse-model (SimpleNegotiator)	1.15	high
Firm hand-built baseline	0.91	medium
Mean over all opponents	1.52	—

[†]few deals, but they emit no usable model of us, so we take the whole Concealing point.

to competitive structure (measured by the correlation of the two utility functions; a domain is labelled cooperative when this correlation is ≥ -0.2), in both move orders. We report our *absolute mean score* (the tournament metric), separating per-opponent results from a cooperative vs. competitive split. We also measure τ_A/τ_B directly, using the same full-outcome-space Kendall computation that the scorer uses.

Headline result. Table 2 summarises performance. Anchor scores positively against *every* opponent type, with a mean score of 1.52. It extracts heavily from conceders (1.6–1.8), takes the whole Concealing point from strong opponents that emit no usable model of us (Tough/MiCRO, ~ 1.0 with a near-zero deal rate), and contests the model-capable agents (BOA/MAP 1.10, Simple 1.15, a firm baseline 0.91). Cooperative and competitive domains score comparably (1.80 vs. 1.48), indicating that the strategy is not tuned to one domain type.

Ablation: what helped, and what did not. Table 3 records the mechanisms we tested. Two changes drove most of the improvement: early-time weighting and the score-first end-game, together with the non-degeneracy fix. The *negative* results are worth as much: every deliberate-deception mechanism we tried was either counter-productive or pointless.

Why deception fails here. The scorer computes Kendall- τ over the *entire* outcome space, not over the offers actually exchanged. A few noisy or decoy offers therefore cannot cheaply corrupt the opponent’s ranking of the whole domain; and any offer we make that deviates from our true preferences either reveals genuine high-utility regions (raising τ_B) or costs us Advantage. The only deception that lowered τ_B at all (decoy-issue freezing against weight-learning models) still cost more Advantage than the shared point it returned. This is direct evidence for the operating point we assumed in Section 1.

Table 3: Ablation. “Kept” mechanisms are in the shipped agent; “Rejected” were implemented, measured, and disabled (left flag-gated for reproducibility).

Mechanism	Measured effect
<i>Kept</i>	
Anti-exploitation concession floor	Removes catastrophic exploitation (vs. a firm baseline: $-0.48 \rightarrow$ competitive).
Early-time weighting (vs. recency)	τ_A $0.48 \rightarrow 0.57$; Concealing share $0.48 \rightarrow 0.50$.
Score-first end-game (close deals)	Firm-opponent score $1.07 \rightarrow 1.19$; mean score up.
Non-degeneracy guard on the model	Mean score $1.41 \rightarrow 1.52$ (recovers Concealing vs. early-accepting opponents).
<i>Rejected (measured counter-productive or neutral)</i>	
Offer-diversification “deception”	<i>Raises τ_B</i> : spreading offers reveals more of our high-utility region.
Decoy-issue freezing	Lowers a weight-learner’s τ_B by 0.02, but costs more Advantage than it returns (vs. a firm baseline, -0.08 score).
Bayesian issue-weight model	Within noise of our frequency model on τ_A .
Runtime opponent-type switching	Marginal ($+0.07$ on four hard-liners only) vs. complexity and gaming risk.
Earlier rescue / selection re-tunes	Flat or net-negative on score.

Conclusions

Anchor treats the ANL 2026 score for what it is: an absolute-score objective in which Advantage dominates and Concealing is a shared point whose modelling half is worth chasing and whose deception half usually is not. Built on an early-weighted frequency model, an anti-exploitation Boulware floor, and a deal-closing end-game, it scores positively against every opponent we tested (mean ≈ 1.52). Two extensions remain open: a late-triggering detector that adapts only the concession exponent to the opponent’s type, and concealment tuned against the specific weight-learning models in the live field, where our decoy experiment shows the mechanism works but did not yet pay on our test set.

References

- [1] Y. Mohammad, S. Nakadai, A. Greenwald. NegMAS: a platform for automated negotiations. *Int. Conf. on Principles and Practice of Multi-Agent Systems (PRIMA)*, 2019.
- [2] T. Baarslag, K. Hindriks, M. Hendriks, A. Dirkzwager, C. Jonker. Decoupling negotiating agents to explore the space of negotiation strategies. *Novel Insights in Agent-based Complex Automated Negotiation*, 2014.
- [3] T. Baarslag, M. Hendriks, K. Hindriks, C. Jonker. Learning about the opponent in automated bilateral negotiation: a comprehensive survey. *Autonomous Agents and Multi-Agent Systems*, 30(5):849–898, 2016.
- [4] ANAC Automated Negotiation League (ANL) 2026. <https://anac.cs.brown.edu/anl>.