

TjAgent - A Hybrid Rule-Based and Probabilistic Bilateral Negotiator

Abstract : Under the Stacked Alternating Offers (SAO) protocol of the Automated Negotiating Agents Competition (ANL 2026), agents must bargain efficiently in environments characterized by incomplete information, strict deadlines, and unpredictable opponent behavior. This report presents TjAgent, a hybrid bilateral negotiating agent that harmonizes a time-dependent deterministic concession curve, a Shannon Entropy-based issue weight estimation model, Laplace frequency smoothing, a tight anticipatory acceptance engine, and low-overhead localized Pareto filtering. Theoretical and empirical evaluation demonstrates that TjAgent exhibits remarkable robustness against both exploitative and stubborn opponents while driving negotiations toward highly efficient Pareto outcomes.

1. Introduction

The fundamental challenge of bilateral negotiation lies in managing the equilibrium between utility maximization and the probability of agreement. Premature concessions deplete the agent's surplus, whereas stubborn stance patterns frequently precipitate negotiation breakdowns.

Designed specifically for the ANL 2026 tournament mechanics, TjAgent is engineered around four core tenets: information concealment, adaptive concession tracking, automated opponent profiling, and proactive stalemate mitigation. Its internal architecture divides these tasks across three tightly coupled modules: a layered, time-sensitive acceptance firewall; a joint Bayesian-Shannon preference learning engine; and a randomized, Softmax-distributed masking bidding strategy.

2. System Architecture & Core Mechanics

2.1 The Baseline Bidding Core (B2 Core)

TjAgent bases its baseline bidding trajectory on a modified Boulware time-dependent concession curve. The target utility $U_{target}(t)$ at relative time $t \in [0,1]$ is formally defined as:

$$U_{target}(t) = U_{max} - (U_{max} - R) \cdot t^{\kappa}$$

Where R denotes the agent's reservation value and κ represents the concession exponent.

To optimize performance across different domains, TjAgent implements an adaptive space scaling approach that dynamically self-calibrates κ based on the cardinality enumerated outcome space size S . If $S \leq 50$, the agent sets $\kappa = 8.0$ (maintaining a stubborn posture early on, followed by rapid late-stage concessions). If $S > 50$, it sets $\kappa = 6.0$ (smoother concession) to guarantee a wide array of exploratory samples for opponent modeling.

2.2 Layered Defensive Acceptance Engine

The acceptance mechanism acts like a tiered institutional firewall, successfully neutralizing late-game predatory behavior or microscopic trap offers through sequential evaluation steps:

① First, an absolute reservation threshold immediately rejects any offer yielding a utility lower than the agent's reservation value R .

② Second, a micro-space corner protection rule blocks irrational or predatory corner offers in ultra-small outcome domains near the absolute deadline ($t \geq 0.995$).

③ Third, an exceptional utility capture rule instantly accepts any incoming offer that meets or exceeds an exceptionally high utility threshold ($t \geq 0.95$).

④ Fourth, a tight anticipatory acceptance mechanism triggers an early agreement if the opponent's current offer yields a utility greater than or equal to the agent's calculated counter-offer for the subsequent round.

⑤ Fifth, a dynamic stalemate floor decay rule scales down the target utility by 15% if a continuous deadlock is diagnosed in mid-to-small spaces. Finally, if the deadline is reached ($t \geq 0.98$) and the offer yields any positive surplus ($U_{offer} > R + 0.005$), it is forcefully accepted to avert a catastrophic break.

2.3 Preference Learning via Shannon Entropy and Laplace Smoothing

To map the opponent's utility structure without explicit preference disclosure, TjAgent implements a frequency-based mapping architecture combined with an Exponential Moving Average (EMA). The model tracks historical issue distributions and utilizes Laplace Smoothing to gracefully account for unseen value combinations:

$$p_i = \frac{c_i + 1.0}{total + |V_i|}$$

By computing the Shannon Entropy of each issue, the agent reasons that lower entropy implies higher opponent rigidity and lower willingness to concede on that specific issue. The agent then inverts this entropy value to derive predictive weights

W_i for each issue:

$$W_i = \frac{1}{Entropy_i + 0.5}$$

The instantaneously deduced utility matrix is fed into a mapping function and smoothed via an exponential moving average ($\alpha = 0.35$) to dampen erratic, non-rational deviations or strategic noise introduced by the opponent.

2.4 Conditional Pareto Filtering & Concealing Softmax Strategy

In the late-game phase ($t > 0.70$) within large spaces ($S > 50$), the agent

activates a conditional Pareto filter. It constructs a candidate set within an immediate target utility band $[U_{target} - 0.005, U_{target} + 0.03]$ and runs a localized non-dominated sorting against its estimated opponent model, effectively filtering out sub-optimal, win-lose choices. To prevent the opponent from reverse-engineering its exact utility curve, TjAgent avoids a deterministic best-pick approach. Instead, it probabilistically samples counter-offers from the filtered candidate pool based on a Softmax distribution scored by a combination of its own utility and the estimated opponent utility, modulated by a time-decaying temperature parameter τ :

$$Score = \frac{U_{own}(c) \cdot [U_{opp_est}(c)]^\gamma}{\tau}$$

Where γ acts as a late-game weight amplifier for the opponent's utility. This obfuscates true reservation bounds while driving offers towards mutual efficiency.

3. Code Sanity & Runtime Optimization Analysis

During Softmax computation, the score term is safely bounded using a clipping safeguard ($\min(score, 500)$) to prevent runtime overflow errors when the temperature

parameter drops near the absolute deadline ($\tau \rightarrow 0.01$). Additionally, the implementation encapsulates the assignment of the opponent utility function within a robust try-except structure, ensuring full functional compatibility across various dependency versions regardless of argument format constraints.

For high-dimensional multi-issue spaces, the explicit $O(N^2)$ double-loop within the Pareto filter can scale poorly if candidate pools swell. Given that outcomes are natively sorted by our own utility in descending order, an optimization utilizing a single-pass linear prune could further minimize the computational footprint under tight step limits.

4. Conclusion

TjAgent successfully balances heuristic stability with adaptive machine learning primitives. Through its layered acceptance matrix, automated deadlock detection loops, and probabilistic concealing bidding strategy, the agent securely protects its utility margins while simultaneously exploring highly cooperative joint gains. The design fulfills all criteria of an elite, robust competitor tailored for the ANL 2026 circuit.