Project 3

CS 332, Fall 2025

Ben Cole & Koshi Harashima
Northwestern University

5 November, 2025

Outline

1. Outcomes from No-regret Learning in Games

2. Manipulability of No-regret Learners in Games

Outline

1. Outcomes from No-regret Learning in Games

2. Manipulability of No-regret Learners in Games

Setting: Repeated First-Price Auction

- Full Information + Full Feedback
- Two bidders; each bidder's value v is fixed across rounds
- FPA; utility if win: $u = v - b$; else 0
- Tie-breaking : if bids are even, allocate a thing randomly

Part 1 Setup

- Rounds and runs: $n \approx 10000$, $n_{mc} = 10$
- Discretization and values: $k = 100$, $v_1 = v_2 = 1.0$
- Algorithms compared
  - **Myopic**: maximizes current-round $\mathbb{E}[u(b)] = (v - b)\Pr(\text{win} \mid b)$ using empirical opponent bids (full feedback), with tie probability 0.5.
  - **Flexible**: exponential weights over k discretized bids; full-feedback updates all arms each round; $\epsilon = \sqrt{\log k / n}$.

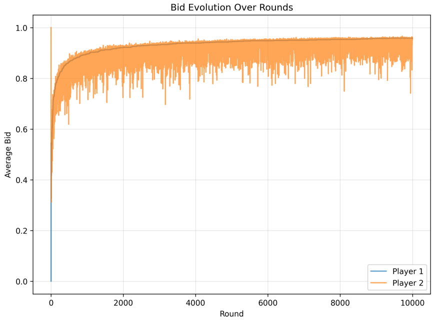
Myopic (one-step expected utility)

- Full-feedback: uses opponent bids from history.
- Discrete bid grid of size k in $[0, v]$; first round bids $\approx 0.5v$.
- For each b on the grid, compute $P(\text{win} \mid b) = \Pr(\text{opp} < b) + 0.5 \Pr(\text{opp} = b)$.
- Choose $b^\star = \arg\max_b (v - b)\, P(\text{win} \mid b)$; ties handled with 0.5.

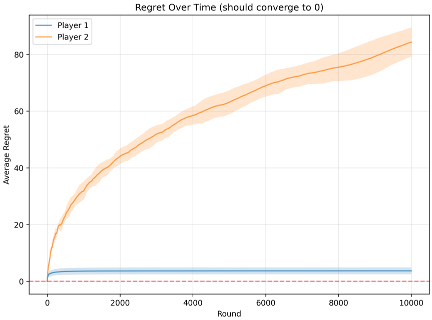
Flexible (Exponential Weights)

- Exponential weights over k bids $b_j \in [0, v]$; full-feedback updates all arms each round.
- Utility per arm: $u_j = (v - b_j) \cdot \Pr(\text{win}_j)$ with $\Pr(\text{win}_j) \in \{0, 0.5, 1\}$ from last opponent bid.
- Cumulative payoffs $V_j \leftarrow V_j + u_j$; sample with $\pi_j \propto (1+\epsilon)^{V_j/h}$, $\epsilon = \sqrt{\log k/n}$.
- First round bids near v on the grid; uses shared grid and tie tolerance.

Part 1 Results: Myopic vs Flexible(OPT lr)



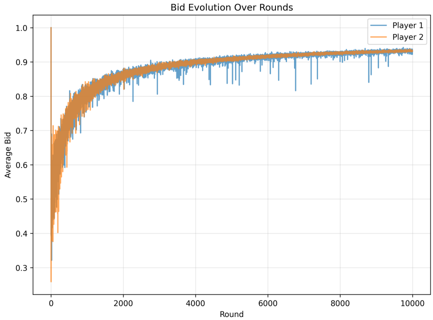
Bid



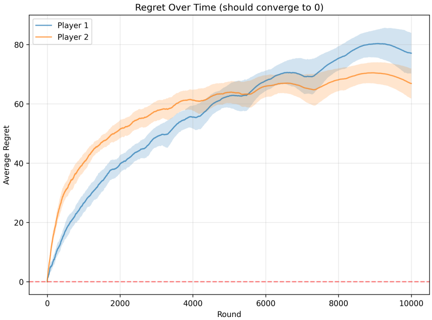
Regret

- Flexible achieves higher average utility and win rate across MC runs
- Regret (vs best fixed bid) is not uniformly lower for Flexible with $k=10$.
- Flexible bids are smoother over time; myopic reacts to recent bids.

Part 1 Results: Flexible(3x OPT lr) vs Flexible(OPT lr)



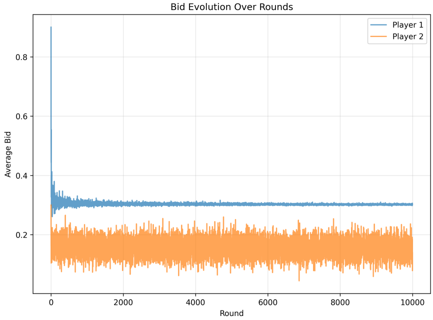
Bid



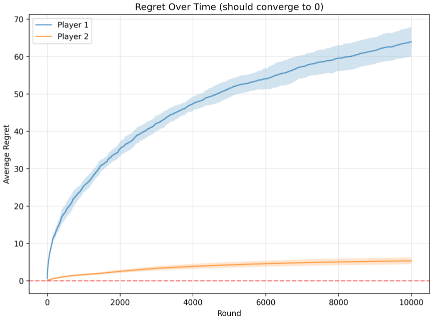
Regret

- speed of adaptation reflects difference of learning rates.

Part 1 Results: Flexible(OPT lr) vs Flexible(OPT lr) with different values



Bid



Regret

- just for comparing to Part 2. If players have different values with similar algorithm, one part always win while the another one always lose.

Robustness

- Ran with $n \in \{100, 1000\}$, $k \in \{10, 100\}$, values $(1, 1)$ and $(12, 8)$.
- Flexible maintains higher average utility and win rate across tested settings.
- Scaling ϵ by $\{0.5, 1, 2\}\times$ default changes variance, not ranking.
- No pure NE in complete-info FPA because both players have incentive to deviate from NE to bid higher as long as below value.

Outline

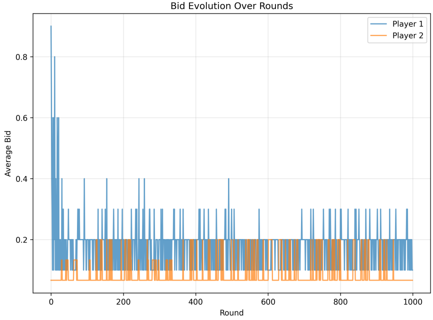
1 Outcomes from No-regret Learning in Games

2 Manipulability of No-regret Learners in Games

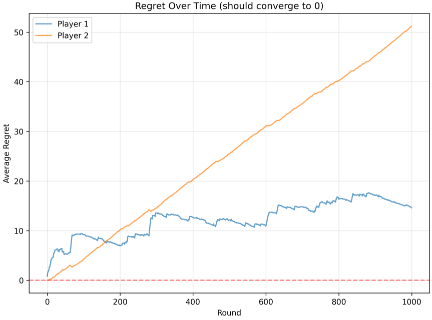
Part 2 Setup

- Rounds and runs: $n = 1000$, $n_{mc} = 1$(for reviewing what is happening!)
- Discretization and values: $k = 10$, $v_1 = 0.9$, $v_2 = 0.3$
- Exploitation observation rounds: 5
- Manipulation Setup:
  - Opponent to be exploited will use Flexible algorithm
  - **Our exploitation**:
    - Observe for initial rounds (e.g., bid $0.2v$) and track opponent bids.
    - Exponential smoothing predicts when opponent drops below a threshold.
    - When predicted low, bid to win with positive margin (handle ties at 0.5).

Part 2 Results: Exploiting Fixed-Lowering Opponent



Bid



Regret

- Sometimes the lower-valued player wins, after which the higher-valued player raises their bid once, but then gradually lowers it again — and this pattern repeats.

Usage of AI

AI was used for figures, code, and design; final review and responsibility by the authors.