Unfair Bidding in Auctions Between LLM Agents

Lior Ben Yaakov
Benyaakov5@mail.tau.ac.il

Roei Ben Zion
Roeibenzion@mail.tau.ac.il

Amnon Levi
amnonlevi@mail.tau.ac.il

Noam Steinmetz Yalon
ns3@mail.tau.ac.il

1 INTRODUCTION

One common type of auction scenario is where several users engage in a repeated online auction, each of them assisted by a learning agent. A typical example is advertisers that compete for ad slots: Typically, each of these advertisers enters his key parameters into some advertiser-facing website, and then this website's "agent" participates on the advertiser's behalf in a sequence of auctions for ad slots. Often, the auction's platform designer provides this agent as its advertiser-facing user interface. In cases where the platform's agent does not optimize sufficiently well for the advertiser (but rather, say, for the auctioneer), one would expect some other company to provide a better (for the advertiser) agent. A typical learning algorithm of this type will have at its core some regret-minimization algorithm, such as "multiplicative weights", "online gradient descent", or some variant of fictitious play, such as "follow the perturbed leader".

A recent paper, Auctions between Regret-Minimizing Agents [7] analyzes the performance of non-human agents based on regret-minimizing algorithms that participate in repeated online auctions on behalf of human bidders. Since this work is a continuation of this paper, it will occasionally be referenced by simply "the paper" or other simple references. The paper provides both simulation results and theoretical analysis for various auction types, such as first-price auctions, second-price auctions, and their generalized variations, while using regret-minimizing agents. The work examines classic regret-minimizing algorithms such as multiplicative weights and follow-the-perturbed-leader.

1.1 Preliminaries

In their experiments, Kolombus & Nisan [7] study the repeated auction setting where the same group of bidding agents repeatedly participate in an auction game, with a fixed auction rule and the same set of items being sold at each round. Each agent's valuation, assigned by its user, remains constant throughout the process, while utilities are additive with items. That is, a bidder with value v for an item, winning k such items, has a total value of kv. Player values and bids are assumed to be on an ϵ-grid, meaning they are multiples of a minimum value $\epsilon > 0$. It is also assumed that agents cannot overbid, and ties are broken uniformly at random.

First-price and second-price auctions follow the standard definitions. For a single item auction, let $b_1 > b_2 > \cdots > b_n$ represent the bids of n bidders. The bidder $s \in [n]$ with the highest bid (b_1) wins the auction and pays a price p. In a first-price auction, the winner pays $p = b_1$, whereas in a second-price auction, the winner pays $p = b_2$. The utility for player s with value v_s is $v_s - p$ if they win the item, and zero otherwise.

We focus on agents using regret-minimization algorithms. The regret of player i at time T, given a bid history $(b^1, \ldots, b^T)$, is defined as the difference between the optimal utility of using a fixed bid in hindsight and the actual utility:

$$R_T^i = \max_b \sum_{t=1}^{T} u_i(b, b_{-i}^t) - u_i(b_i^t, b_{-i}^t),$$

where b_i^t is player i's bid at time t, and b_{-i}^t represents the bids of the other players at time t. A regret-minimizing agent ensures that $R_T^i/T \to 0$ almost surely as $T \to \infty$.

We use the following notation to describe the empirical dynamics of bids. Let Δ represent the space of probability distributions over bid tuples in an auction with discrete bid levels. The empirical distribution of bids after t rounds in a sequence of $T \geq t$ auctions is denoted by $p_T^t \in \Delta$, and $p_T^t(b)$ is the empirical frequency of a bid tuple b after round t. We define convergence of dynamics as follows:

Definition 1. *The dynamics converge to a distribution $p \in \Delta$ if for every $\epsilon > 0$, there exists $T_0(\epsilon)$ such that for every $T > T_0$, with probability at least $1 - \epsilon$, for every $\epsilon T < t \leq T$, $|p_T^t - p| < \epsilon$.*

Intuitively, this implies that after a sufficiently long time, the time-averaged distribution stabilizes near a stationary distribution. The authors further note that convergence of the cumulative empirical distribution to a joint distribution does not imply the convergence of any single agent's play to a distribution over its own actions, as studied in other works.

1.2 Co-Undominated Coarse Equilibria

First, the definition of weak domination:

Definition 2. *Let A, B denote subsets of the action spaces of the two players in a two-player game. Action $i \in A$ of player 1 is called weakly-dominated-in-B by action i' of player 1 if: $\forall j \in B : u_1(i, j) \leq u_1(i', j)$ and $\exists j \in B : u_1(i, j) < u_1(i', j)$.*

Next, co-undominated CCEs, aimed to capture CCEs in which players use only strategies that are not weakly dominated (relative to the distribution of the other players):

Definition 3. *Let $p_{i,j}$ be a CCE of a (finite) two-player game with action spaces I, J. Denote its support (A, B) by $A = \{i \in I \mid \exists j \in J \text{ with } p_{i,j} > 0\}$ and $B = \{j \in J \mid \exists i \in I \text{ with } p_{i,j} > 0\}$. A CCE is called co-undominated if no strategy in its support is weakly dominated relative to the support of the other player, i.e., if for every $i \in A$ and every $i' \in I$, action i is not weakly-dominated-in-B by i', and similarly for B.*

Finally, [7] provides a proof to the following lemma connecting mean-based regret algorithms with the newly defined co-undominated CCEs:

Lemma 1. *Consider the dynamics of any mean-based regret-minimizing agents playing a repeated finite two-player game. If the dynamics converge to a distribution p, then p is a co-undominated coarse correlated equilibrium.*

1.3 Second-Price Auction

The paper provides an analysis of a second-price auction involving two bidders and a single item. In standard setting this type of auction is known to be dominant-strategy incentive compatible (DSIC), meaning bidders have no incentive to misreport their values. Consequently, the bidder with the higher value wins and pays the amount equal to the lower bidder's value. However, the paper demonstrates that in auctions between two multiplicative-weights agents, in the limit, the agent that was given the higher valuation wins but, surprisingly, pays an average price that is strictly less than the second price. Formally, they prove the following Theorem:

Theorem 1. *In the repeated second-price auction with discrete bid levels that are multiples of ε and players using multiplicative-weights algorithms with values $v > w$, in the limit the dynamics converge to a joint distribution in which for the "high player" with value v, $Pr[0] = Pr[\varepsilon] = \ldots = Pr[w] = 0$ and $Pr[w+\varepsilon] = Pr[w+2\varepsilon] = \ldots = Pr[v] = \frac{\varepsilon}{v-w}$, and for the "low player" with value w, $0 < Pr[0] \le Pr[\varepsilon] \le \ldots \le Pr[w]$.*

In addition to proving Theorem 1, the authors execute simulations that show consistent results. Specifically, for two agents with values $v = 1$ and $w = 0.5$, the long-term bid distribution of the "high player" is uniform in $(0.5, 1]$, while the "low player's" bid distribution has full support on $[0, 0.5]$. Moreover, the "high player" always wins (except for a small number of times at the start), and pays an average price of 0.27.

The discrepancy in the average price paid suggests that a bidder with a valuation that is higher than the average paid price and lower than the highest valuation can benefit from misreporting their values. For example, a bidder with a value of $v = 0.4$ bidding against a bidder with a value of $w = 0.5$ could gain utility by reporting his value as $v = 1$, leading to the simulated dynamics. This would ensure his win and positive utility of $u = 0.4 - 0.27 = 0.13$. Hence, despite the agents partaking in a second-price auction, the humans do not observe it as such, and they may be able to improve their utility by misreporting their value. The authors further show that a similar phenomenon occurs in generalized second price.

Considering the above, given full information, the best reply of the high player is always to bid as high as possible. However, if the value of the other player is not known, there exists multiple possible equilibria of the meta-game, and it is unknown how to select one as the outcome of specific user behavior. Likewise, how to give a reasonable prediction of the revenue in equilibrium also remains unknown.

1.4 First-Price Auction

When analyzing first-price auctions, the authors first start by looking at a subclass of an agent-based auction, where the agents are both regret-minimizing and mean-based. The following theorem is proven for this scenario:

Theorem 2. *In a repeated first-price auction between two mean-based regret-minimizing agents with possible bid levels that are a discrete ϵ-grid, If the dynamics converge to any single distribution, then the high player wins with probability approaching 1 and pays an average price that converges to the second price (up to $O(\epsilon)$).*

In simpler terms, for this scenario, as ϵ gets smaller (relative to the difference between the player's values), the agents converge to a second-price auction over the bidders' reported values. As stated previously, a direct second-price auction is DISC, and thus it means that this subclass of agent-based first-price auction is DISC as well. This is in contrast to the case of direct bidding, where the first-price auction is known to not be DISC since users can benefit from misreporting their true value.

It is important to note that the theorem *does not* prove that the dynamics converge, but rather assumes it. However, based on their simulations of agents using the multiplicative-weights algorithm, the authors conjecture that this assumption is irrelevant.

Next, the authors analyze a generalized first-price auction in a specific game: a first-price two-slot "ads" auction, with the top ad slot having a click-through rate (CTR) of 1, and the bottom slot yielding a CTR of 0.5. The winner takes the top slot and gets their value minus their paid bid, while the losing player takes the bottom slot and gets half of their value. The following mixed Nash equilibrium is proven:

Theorem 3. *Assume w.l.o.g. that $v \ge w$. In the unique Nash equilibrium of the two-slot auction described above, the players mix their bids x, y, respectively, according to the following cumulative density functions with support $[0, w/2]$.*

$$F(x) = \frac{x}{w-x} \qquad G(y) = \left(s\frac{v}{w} - 1\right)\frac{y}{v-y}.$$

The expected payoffs are $u_1 = v/2 + (v-w)(1 - ln(2))$, and $u_2 = w/2$.

While this Nash equilibrium does emerge from the theoretical analysis, the simulations of a symmetric game using the multiplicative-weights algorithm do not converge to it and only reach coarse correlated equilibrium (CCE). The following analysis is provided for the predicted (and empirically observed) CCE.

First, note that there is no point in bidding higher than half of your value since winning with such a bid would yield less than losing, which yields half of your value. Denote v the symmetric value for both players. Given a bid x, if $x + \epsilon$ is smaller than $v/2$, then that is the best reply to x. Otherwise, 0 becomes the best reply (since losing the auction is better in this case). This analysis predicts a dynamic of gradually increasing bids until a mutual drop to 0 upon reaching $v/2$. This is backed by the observed simulations.

The paper contains only experimental results for non-symmetric auctions and the authors leave their theoretical analysis as a future work, providing only an unjustified prediction.

Finally, the authors analyze the meta-game played by the symmetric users deploying the agents. They conclude that an ϵ-Nash equilibrium is achieved when one player reports an arbitrarily high value v and the other player reports $w = 1/(6 \cdot (1 - ln(2)) \approx 0.54$ (of a true value of 1), where $\epsilon = O(1/v)$. The respective utilities are $5/6 - O(1/v)$ and $1/2 + O(1/v)$. If there exists a maximum allowed bid, then this becomes an exact equilibrium. This theoretical result is further supported by experimental simulations.

An interesting corollary from this result is that in these equilibrium points, the users capture 8/9 of the welfare, exhibiting (a probably unwanted) implied collusion.

2 RELATED WORK

Regret Minimization. Is a strategy used in decision-making where the goal is to make choices that minimize regret, which is the difference between the actual outcome and the best possible outcome in hindsight. More formally, given the horizon $T > 0$, a trajectory of actions $a_t, t \in [T]$ and the optimal action in each timestep $a_t^*, t \in [T]$, the regret $R(T)$ is defined as:

$$R(T) := \sum_{t=1}^{T} (a_t^* - a_t)$$

In online learning or reinforcement learning, it focuses on selecting actions that, over time, minimize the accumulated regret compared to the optimal strategy, improving performance progressively. Some notable algorithms for regret minimization include the Multiplicative Weights (MW) algorithm, in which we learn a probability of actions, and the expected regret is $O(\sqrt{\frac{ln(n)}{T}})$ (where n is the number of actions). Other approaches include fictitious play [2] ("follow the leader", FTL), which is a simple strategy used in online learning where, at each round, the algorithm chooses the action that would have performed best based on past data. Other lines of work assume various different settings and include (among others) Multi-Armed-Badits, and variants of MW (e.g. EXP3).

Auction Theory. The classic first/second price auctions have been studied extensively [9]. A multiple-item auction includes many real-world problems, such as ad auctions, where there are multiple time slots to be distributed to bidders with ads. Another auction setting is the keyword matching setting, used by search engines. In this setting, the bidders compete over the engine preference given a keyword. Since these auctions happen in microseconds, humans are replaced by automated agents (auto-bidding). This is an example of a real-life auction setting that is similar to the auctions discussed in the paper.

Meta games. Coined in [8], the meta game model refers to a situation where the human users "instruct" their learning automated agents using parameters, and the agents play behalf of their users. The auto-bidding setting described above is a concrete example of it. The main question asked in the paper is whether the human users should provide the agents with their "true valuation", or should they misreport in order to increase their utility.

LLMs in auctions. The advent of LLMs reasoning skills has been a motivation for incorporating them in auctions. One prior work [5] discussed the idea of auctioning for a place in a summary generated by an LLM that is intended to provide the user with insightful products. Other work [3] has created a simulation framework that involves LLMs in an auction setting in order to enhance planning and execution. However, this was not evaluated in the auto-bidding setting where the agents are the bidders rather than the actual players. We intend to evaluate the role LLMs can play in the auto-bidding setting, and check if the phenomenons observed in the last paper still hold when LLMs are involved.

3 MAIN RESEARCH QUESTION

Upon reviewing the paper, we identified several potential directions for extending its research and findings. These include exploring agents that do not exclusively or primarily minimize regret, analyzing scenarios with more than two bidders, testing across a broader range of auction types or alternative "games" and assigning multiple agents to each user to introduce internal competition that could enhance performance for the bidders, among other possibilities.

One notably promising approach involves replacing the multiplicative weights (MW) algorithm with a deep neural network to evaluate its behavior, performance, and impact on the paper's conclusions. This direction is particularly compelling given the remarkable performance of deep neural networks (NNs) in recent years, often exceeding the results achieved by "traditional" algorithmic approaches across a variety of tasks. We aimed to evaluate the performance of basic auction formats when agents are powered by neural networks and to determine whether the results from the original paper replicate in this setting.

4 METHODS

As described before, we opted to replace the MW algorithm with a deep neural network, specifically, a large language model. Concretely, the large language model (LLM) is prompted with the bidding and reward history to generate the next bid at each step. After iterative experimentation, we selected the LLaMA 3.2 model [4] with 3 billion parameters.

The decision to use a large language model as our algorithm, and specifically the LLaMA model, was driven by the following reasons:

(1) In recent years, large language models (LLMs) have emerged as state-of-the-art systems, showcasing exceptional performance after training [1, 4, 10].
(2) We sought an algorithm that, while potentially requiring some pre-training, would impose minimal computational demands in terms of processing time and memory requirements during runtime.
(3) Similar studies [6] have demonstrated that using pre-trained models significantly improves results, strongly motivating us to adopt this approach.
(4) Neural networks (NNs) have shown remarkable success in analyzing time series data and forecasting future outcomes. Although a second-price repeated auction does not strictly qualify as a time series, and large language models (LLMs) are not inherently optimized for time series tasks, there are notable similarities. This makes exploring this approach particularly intriguing.
(5) The LLaMA model is readily accessible through pre-trained implementations available in Hugging Face libraries [11], making it convenient for immediate use and analysis.

Once we selected our algorithm, we faced several key decisions related to the auction process, including defining the most effective prompt for the LLM. The final prompt consistently begins with a system message containing an explanation of the auction process:

`You are a bidder in a multi-round {auction_type}-price auction, meaning that each round: a. The winner is the highest bidder and they pay the {auction_exp} bid and get a reward of valuation - paid amount. b. The loser is the lower bidder and they get a reward of 0. Your valuation is {valuation}. Each round, you must choose a bid only from 0 to {valuation}. In your choice, consider all previous rounds results. You must only ever reply with a single number. Your goal is to maximize your reward.`

We substitute `{auction_type}` with `first/second` accordingly, `{auction_exp}` with `higher/lower`, and `{valuation}` with the bidder's value.

We elected to use unsupervised few-shot learning. To facilitate it, after each round, the prompt is updated by appending the following user message:

`In the previous round, you bid {last_bid} against the other agent's {other_last_bid}. You bid {rank} and {result} that round, getting a reward of {reward}. Bidding {order} would have yielded {other_result} reward. What should your next bid be?`

Here, we replace `{last_bid}` with the agent's previous bid, `{other_last_bid}` with the previous bid of the other agent, `{rank}` with `higher/lower`, `{result}` with `won/lost`, `{reward}` with the received reward, `{order}` with `below/above`, and `{other_result}` with the other possible result.

The final input given to the LLM at round T is

$$I_T = S_0 \parallel \|_{t=T-\mathcal{K}}^{T} (A_{t-1} \parallel U_t)$$

where S_0 is the system message containing the rules, A_{t-1} is the agent's reply at round $t-1$, and U_t is the user message at round t. $\mathcal{K}$ is a hyper-parameter that sets the size of the message history given to the LLM; an extended discussion of it can be found in section 6.

5 EXPERIMENTS AND RESULTS

We run multiple first and second price auctions, each time substituting at least one MW-based agent with an LLM. All of the behaviors described below were consistently observed. In cases where multiple prominent behaviors emerged, all were documented. However, we excluded instances where the LLM appeared to disregard fundamental rules and generated 'illegal' responses. We believe these anomalies may result from the model's limited size and could potentially be addressed through fine-tuning.

5.1 First-price auction

Figure 1 shows the dynamics of first-price auction where one agent is an LLM with a value of 1 and the second agent is MW-based agent with a value of 0.5. We observe similar results to those reported in the original paper: the higher bidder wins the auction and pays a price slightly above the second-highest bid. However, in our case, we replace the higher bidder with an LLM. Despite this modification, the findings remain consistent with the paper, where we get the second-price outcome from the point of view of the bidders, therefore it is still a dominant strategy for all participants to truthfully report their valuations to their respective agents.

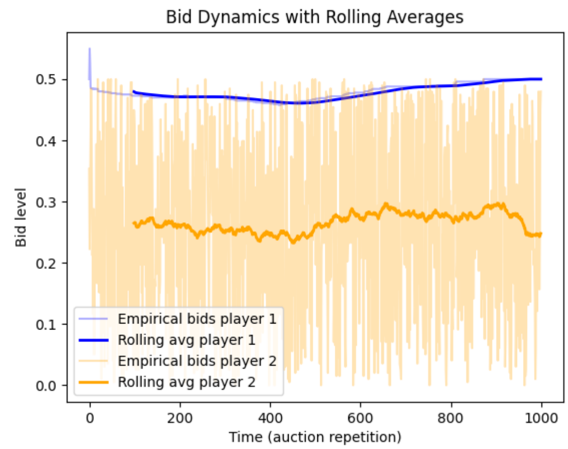


Figure 1: Dynamics of a first-price auction where agent 1 is an LLM with value 1 and agent 2 is a MW-based agent with value 0.5.

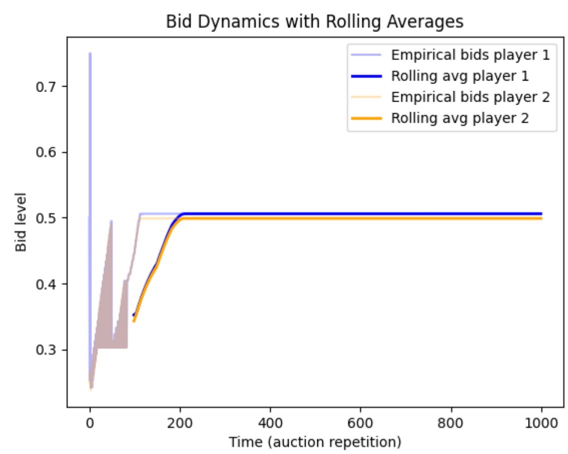
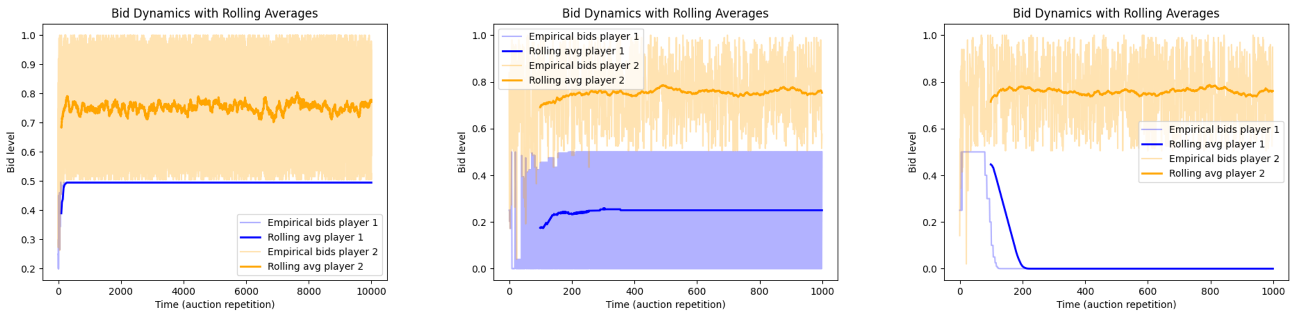


Figure 2: Dynamics of a first-price auction where agent 1 is an LLM with value 1 and agent 2 is an LLM with value 0.5.

Figure 2 shows the dynamics of first-price auction where both agents are LLMs, with values 1 and 0.5. Once again, we observe the same behavior: the higher bidder, which is an LLM, wins the auction and pays a price slightly above the second-highest bid. This suggests that truthfully reporting valuations to the agents remains a dominant strategy. Notably, in this iteration, the second-highest bidder, also represented by an LLM, converges to bidding its true valuation, contrasting with the MW agent's behavior observed both in the original paper and our prior experiment.



(a) **(b)** **(c)**

Figure 3: Dynamics of a second-price auction where agent 1 is an LLM with value 0.5 and agent 2 is a MW-based agent with value 1.

5.2 Second-price auction

Figure 3 shows the dynamics of second-price auction where one agent is an LLM with value 0.5 and the second agent is MW based with a value of 1. We observe that the MW-based agent with the higher valuation consistently bids below the reported value and always wins in the limit. Regarding the price paid, figure 3a demonstrates that the LLM agent with the lower valuation converges to bidding exactly the value reported to it, resulting in the winner paying the 'classic' second price. In contrast, figures 3b and 3c show the LLM agent bidding, on average, approximately half of its valuation or zero, respectively, allowing the winner to pay strictly less than the second price. These scenarios mirror the behavior described in the original paper, where the lower-valuation player could achieve a strictly positive utility on average by misreporting his true value to his agent. Furthermore, in 3b we see convergence to the same value that was reported in the original paper. These three distinct behaviors were repeatedly observed in this setting. This behavioral uncertainty appears to stem from the random sampling employed by the LLM (we kept the Llama model's default temperature setting unchanged).

Figure 4 shows the dynamics of a second-price auction where one agent is an LLM with a value of 1 and the second agent is MW-based with a value of 0.5. We observe the same behavior as described in the paper and in the previous example: the agent with the lower valuation, now an MW-based agent, bids on average approximately half its valuation, while the agent with the higher valuation, now an LLM, bids below its valuation and consistently wins. Once again, this scenario parallels the one in the paper, where the lower-valuation player could benefit from misreporting his true value to his agent.

Lastly, figure 5 shows the dynamics of second-price auction where both agents are LLMs, with values 1 and 0.5. We observe an intriguing behavior wherein, after a significant number of rounds, the bidding stabilizes with both agents consistently placing extremely low bids and alternating wins. This bidding pattern leaves minimal revenue for the auctioneer and resembles collusive behavior, seemingly working against the interests of the platform designer.

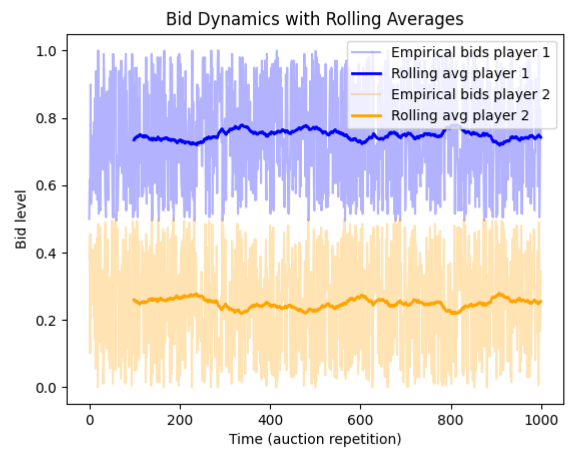


Figure 4: Dynamics of a second-price auction where agent 1 is an LLM with value 1 and agent 2 is a MW-based agent with value 0.5.

6 LIMITATIONS

After selecting our preferred algorithm and defining the auction framework, we encountered several challenges that required careful consideration.

One challenge was determining the appropriate value of $\mathcal{K}$, the amount of historical data provided to the LLM. This decision posed both algorithmic and computational difficulties, as providing a large number of time steps would significantly increase memory usage. To address this, we limited the context provided to the LLM to the last 20–40 rounds, depending on the specific setting.

Second, due to our limited computational resources, we were forced to make additional compromises. Notably, we omitted fine-tuning the model on our specific task and only used few-shot learning. We were also unable to conduct extensive comparisons across different prompt designs.

Third, the MW algorithm updates weights iteratively, starting with randomized values. We faced a critical decision: should the

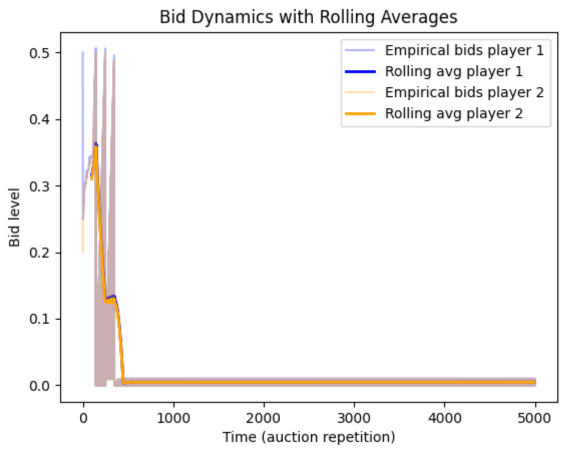


Figure 5: Dynamics of a second-price auction where agent 1 is an LLM with value 1 and agent 2 is an LLM with value 0.5.

LLM predict the next weight adjustment or directly generate the next bid? Ultimately, we opted for the latter approach, though we didn't fully exhaust the exploration of the former.

This series of constraints and decisions shaped the overall structure of our implementation and limited the extent to which we could optimize our approach.

7 DISCUSSION

In this research, we conducted first-price and second-price auction experiments, substituting at least one MW-based agent with an LLM. Our key findings include that, in first-price auctions, an LLM as the higher bidder consistently won and paid prices resembling second-price outcomes, reinforcing truthful reporting as a dominant strategy, consistent with the original paper. In second-price auctions, an LLM as the lower bidder either bid its exact valuations, resulting in classic second-price outcomes, or below it, enabling the winner to pay less. This latter behavior aligns with the original paper's finding that lower-valuation players can benefit by misreporting their values.

Overall, the empirical results from the original paper were largely reproduced. While theoretical analysis of LLMs remains highly challenging, our findings suggest that some theoretical results from the paper may extend to these settings as well.

Future work could explore several promising directions to enhance and extend the current study. One avenue involves improving the LLM model itself by extending the context window, increasing the number of parameters, performing fine-tuning for auction-specific tasks, or training a new model from scratch using a different architecture. Optimizing prompt design also offers a valuable research direction. For example, studying how different input representations — such as natural language descriptions of valuations versus numerical inputs — affect agent strategies and performance. It would be interesting to see what will be the impact of those changes on agent performance and adherence to auction rules.

Another important focus could be further investigating agent strategies, particularly examining whether LLMs exhibit learning, collusion, and their ability to adapt and exploit their opponents' strategies. It would be interesting to identify the factors that influence the emergence and conversion to specific behaviors.

Additional studies could test LLMs in more complex auction formats, such as GSP or GFP, or expand experiments to include a larger number of bidders. Research into generalization could examine whether LLMs trained on one auction format effectively transfer knowledge to others, leveraging pre-training or transfer learning to enhance their adaptability to new tasks.

Lastly, testing LLM-powered agents in real-world auction platforms could provide insights into their practical feasibility and performance, bridging the gap between experimental settings and applied environments.

8 ACKNOWLEDGMENTS

We thank Noam Nisan and Yoav Kolumbus for the insightful discussions.

REFERENCES

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[2] C. F. Carter. 1952. *The Economic Journal* 62, 247 (1952), 625–628. http://www.jstor.org/stable/2226909

[3] Jiangjie Chen, Siyu Yuan, Rong Ye, Bodhisattwa Prasad Majumder, and Kyle Richardson. 2024. Put Your Money Where Your Mouth Is: Evaluating Strategic Planning and Execution of LLM Agents in an Auction Arena. arXiv:2310.05746 [cs.CL] https://arxiv.org/abs/2310.05746

[4] Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783* (2024).

[5] Kumar Avinava Dubey, Zhe Feng, Rahul Kidambi, Aranyak Mehta, and Di Wang. 2024. Auctions with LLM Summaries. arXiv:2404.08126 [cs.GT] https://arxiv.org/abs/2404.08126

[6] Zhe Feng, Guru Guruganesh, Christopher Liaw, Aranyak Mehta, and Abhishek Sethi. 2021. Convergence analysis of no-regret bidding algorithms in repeated auctions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 5399–5406.

[7] Yoav Kolumbus and Noam Nisan. 2022. Auctions between Regret-Minimizing Agents. In *Proceedings of the ACM Web Conference 2022* (Virtual Event, Lyon, France) *(WWW '22)*. Association for Computing Machinery, New York, NY, USA, 100–111. https://doi.org/10.1145/3485447.3512055

[8] Yoav Kolumbus and Noam Nisan. 2022. How and Why to Manipulate Your Own Agent: On the Incentives of Users of Learning Agents. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 28080–28094. https://proceedings.neurips.cc/paper_files/paper/2022/file/b39fcf2e88dad4c38386b3af6edf88c7-Paper-Conference.pdf

[9] Vijay Krishna. 2003. *Auction Theory*. Elsevier Inc., United States. https://doi.org/10.1016/B978-0-12-426297-3.X5026-7

[10] Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien Vincent, Zhufeng Pan, Shibo Wang, et al. 2024. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *arXiv preprint arXiv:2403.05530* (2024).

[11] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*. 38–45.