

Note On "Mechanism Design for Large Language Models"

Paul Duetting, Vahab Mirrokni, Renato Paes Leme, Haifeng Xu, and Song Zuo

June 1, 2024

1 Introduction

For application in ads, consider making LLM produce joint ads for multiple products.

Challenges

- **Modelling and Expressing Preferences:** In auction models, there usually exist value functions that can assign a value to each outcome. However, as generative models, LLMs don't attribute values to each example, but rather succinctly encode preferences over outcomes in a stateless neural network model that predicts continuation probabilities.
- **Necessity of Randomization:** LLMs typically have a worse performance, when forced to output deterministically, than if they are allowed to sample from a distribution. Therefore, an auction that aggregates LLM outputs must also output distributions.
- **Technical Alignment:** Auction solutions should be technically aligned with current LLM technology.
- **Computational Efficiency:** LLM models are expensive to query, so the auction computation should not add too much overhead.

Contributions

The token Auction Model. The first contribution is a formalism("The Token Auction Model") for studying this problem. The *token auction* this paper proposes operates on a token-by-token basis, and serves to aggregate several LLMs to generate a joint output.

Simple and Robust Token Auctions: In order to design *robust token auctions*, formulate two desirable incentive properties, which this paper consider minimal requirements:

- *Payment monotonicity:* Given two different bids for the same agent, a final distribution is closer to the desired distribution if and only if the payment is higher.
- *Consistent aggregation:* If for two different bids of the same agent, the final distribution is closer to the preferred distribution for some bids of the other agents, then it should be so for all bids of the other agents.

2 Main Body

3 Conclusion