CS908 Assignment 4 - Research Proposal

Thomas Archbold

1 Introduction

Prediction markets are exchange-traded markets[1] that trade on the outcome of events, as opposed to traditional financial instruments. Since actors in the market participate by putting up their own money in the form of betting on an unknown future outcome, the market prices can indicate the beliefs held by the market of certain events occurring. What can make these markets even more interesting is the ability to combine these bids on events in complex ways, giving users the freedom to make predictions over a range of different unknown outcomes. This is the approach taken by *Predictalot* [1], which boasts a platform on which one can bet on over 9.2 quintillion outcomes in the NCAA Men's College Basketball playoffs. This specification outlines the plans to implement a similar combinatorial prediction market for betting on outcomes in the 2020 United States presidential election, placing as few restrictions as possible on the types of bets available to traders.

The rest of the specification is structured as follows. In Section 2 we outline in greater detail the motivation for such a project and present our specific goals for the platform. In Section 3 we discuss the underlying theory on which we will model the platform, including recent results in mechanism design for combinatorial auctions. We also discuss the current software that exists in use by others to implement other prediction markets, as well as how this project will fit in with the current body of work. In Section 4 we outline the methods and approaches we plan to take in order to complete this project, including a brief discussion of the tools, technology, and data that we plan to use. Aspects related to the effective management of the project are discussed in Section 5. We conclude with a brief discussion of progress made thus far in Section 6, followed by closing thoughts and plans for the next stages in Section 7.

2 Project aims

2.1 Overview

Prediction markets provide ways in which to bet on the occurrence of events in the future, and are often used to bet on a variety of circumstances – this could be on the outcomes of a political election, sporting events, or any other probabilistic event. Since there is an incentive to do well in such a market, by players staking their own money or units from some point system, people are inclined to bet how they truly feel about certain events, and hence public sentiment on these events can be crowdsourced to learn how likely they are thought to occur. Combinatorial prediction markets, taking inspiration from the theoretical economics and mechanism design literature, allows for bets to be bought and sold, and it is how these are combined in complex ways that we may learn the most about how the public view the likelihood of separate events as being related. With this project, we hence seek to implement a combinatorial prediction market for the 2020 US presidential election. There is significant global importance to the outcome of this and the individual events that comprise the presidential campaigns that it would provide useful insight into public sentiment on the matter – or at least, more so than for men's college basketball.

It is well-known, however, that computing allocations of goods to buyers maximising, for example, social welfare or revenue, requires solving an NP-hard optimisation problem [7]. Furthermore, for a market offering bets on m separate events, there are 2^m possible ways of combining such bets – how can we expect users to enter an exponential number of bids before they even get to participate in the market? These two problems are the focus of much of the literature within algorithmic mechanism design [6], a subfield within algorithmic game theory. Broadly, this is an area of research at the intersection of economics and computer science that is concerned with designing the ways in which self-interested agents act within a strategic environment and achieving certain economic properties – truthfulness, budget balance, individual rationality, for example – while ensuring that the mechanisms remain practical to

[1] A market in which all transactions are routed through a central source.

implement. It hence makes extensive use of techniques favoured in computer science, most notably asymptotic analysis, randomisation, and approximation. The goal in the modern literature therefore departs from striving to compute the idealistic but impractical optimal solution, towards computing ones which are approximately-optimal, or "good enough".

Even restricting ourselves to approximately-optimal solutions, we are faced with yet another problem – how do we define "good enough"? What makes, say, a 4-approximation for computing an allocation in a two-sided market any better than a 7-approximation that also achieves group-strategy proofness?[2] The answer is that it depends – trade offs must be made on a variety of assumptions on the structure of the market and its agents, and these parameters may be tweaked depending on the setting. Each of these decisions will lead to different mechanisms with different performance. Given that there is no single metric by which we can judge a mechanism's performance, this project also aims to implement software that is generalised enough such that it is capable of tweaking the parameters on the market to see what, if any, impact they have on the functioning of the system as a whole.

2.2 Features

As stated, the main aim of this project is to create a combinatorial prediction market for betting on outcomes in the forthcoming US presidential election. The underlying exchange will be modelled as variants of two-sided combinatorial markets (see Section 3 for a more in-depth discussion of what this entails). To get a concrete idea of the features, we detail them below under the categories of **core**, **additional**, and **stretch** features. The core features will implement all of the functionality to qualify the software as a combinatorial prediction market, and will cover making and buying basic bets, combining them in complex ways, and calculating the somewhat accurate based on the actions of the market, likely based on dummy data for testing; the additional features will extend this functionality so we may begin to explore the effects of making use of different underlying mechanisms from the literature; and the stretch features will focus on transitioning from the dummy data into the real world, as well as the provision of more complex bets and further refinement of the user experience.

2.2.1 Core features

- ass

2.2.2 Additional features

- Two-sided exchange

2.2.3 Stretch features

- ass

3 Background

3.1 Combinatorial auctions

In a single-item auction, we have a collection of n agents who each possess a valuation for acquiring a single item on offer. A *mechanism* is any method for allocating this item to the bidders – typically, this involves specifying an *allocation rule* and a *payment rule*. The former may be responsible for collecting information from the participants, often a bid, to compute the allocation, while the latter is used to ensure that agents act truthfully.

We use the real number v_i to denote agent i's valuation for acquiring the item. This information is private to each agent (referred to as "bidders" or "buyers"), meaning neither the other agents nor the mechanism itself has any way of determining this value. Naturally, we typically wish to allocate this item to the bidder who truly values it the most – that is, allocate it to agent $i^* = \arg\max_{i\in[n]} v_i$.[3] It is a typical for an auction to elicit a bid b_i from each agent so as to acquire their valuation v_i – note, however, that since the agents are strategic they may lie (i.e., submit bid $b_i \neq v_i$) if they believe it is in their best interests. The Vickrey mechanism [8] achieves such an allocation by believing the bidders, giving the item to the agent with the highest bid, and charging him the next highest bid.

[2] These are just examples to illustrate the point – mechanisms achieving these guarantees may or may not exist.
[3] We use the standard notation of $[k]$ to represent the set $\{1, \ldots, k\}$

It is easy to see how this can be extended to the case of multi-item auctions, where we have m items to allocate. Buyers now have a collection of valuation function, $v_i(S)$ that associate to each subset of items $S \subseteq [m]$ a value for acquiring this subset. Here we can see where sources of complexity are introduced into the model – since the number of subsets of any set of size m is 2^m, we cannot hope to either collect all these valuations from bidders nor evaluate them in sub-exponential time, without imposing further restrictions. In the literature, one common way to deal with this obstacle is by restricting the bidders to be k-minded – each bidder submits a maximum of k bids for k different subsets, and effectively a bid of zero for every other set. There are also subtleties involved in allowing items to be homogeneous, in which case the auction contains m copies of the same item, or heterogeneous, where each item is unique. The idea of this project is to model bets on future outcomes as items, and since we will be combining different bets to learn about public sentiment on these outcomes, we assume the items in the auction to be heterogeneous. Furthermore, since we are keen to explore the practical performance of some of the mechanisms in the literature, we will often use the k-minded bidders model in our underlying representation of the auction mechanism.

In a combinatorial auction, instead of simply maximising over n numbers as in a single-item setting, i.e. selecting the single bidder with the highest bid, we must maximise over all possible allocations, wherein lies another source of complexity. The multi-item auction analogue to awarding the item to the agent who values it most is the concept of maximising social welfare, which we denote W. If we compute allocation $A = (A_1, \dots, A_n)$, where agent i is allocated subset A_i, then the value we are maximising as a function of this allocation is $W(A) = \sum_{i \in [n]} v_i(A_i)$. The analogue to the Vickrey auction in the multi-item setting, the Vickrey-Clarke-Groves mechanism [8, 2, 5] again believes the bidders and awards subsets to agents who value them most and charges each bidder their *externality*, or the extra cost in social welfare incurred by that agent participating in the auction.

Up until now we have assumed that the auction itself holds the items – this will not be the case in our prediction market since the agents themselves will make and hence hold the bets. It is therefore useful to have the concept of a *two-sided* market. This comprises a set of two distinct types of agents: sellers, who initially hold the items for sale, and buyers, who are interested in buying the items from the sellers. Using the model of Colini-Baldeschi et al. [3], formally a two-sided market is a tuple (n, m, k, I, G, F), where $[n]$ is the set of buyers, $[m]$ is the set of sellers, $[k]$ is the set of items, and $I = (I_1, \dots, I_m)$ is the *initial endowment*, such that I_j is the set of items initially held by seller j. Vectors $G = (G_1, \dots, G_n)$ and $F = (F_1, \dots, F_m)$ are the distributions from which the buyers' are sellers' valuation functions are assumed to be drawn, respectively. The notion of an allocation changes only slightly under this model: given a two-sided market, our aim is to redistribute the items among the agents so as to maximise the social welfare. An allocation is hence a pair of vectors $(X, Y) = ((X_1, \dots, X_n), (Y_1, \dots, Y_m))$ such that the union of X and Y is the set of items $[k]$, and $X_1, \dots, X_n, Y_1, \dots, Y_m$ are mutually disjoint, meaning no two agents are allocated the same item. As before, buyers have a valuation $v_i(S)$ for each subset of items S, and we introduce the same idea for the sellers, which we denote $w_j(S)$. Our goal is still to maximise the social welfare, which we now write as:

$$W(X, Y) = \sum_{i \in [n]} v_i(X_i) + \sum_{j \in [m]} w_j(Y_j)$$

We conclude our discussion of the underlying by introducing some of the main economic properties pursued in the field of mechanism design, useful in our discussion of the current literature:

- **Incentive Compatibility (IC):** agents are incentivised to bid truthfully (i.e. submitting $b_i = v_i$) as they can do no better by lying – truth-telling is a *dominant strategy.*
- **Individual Rationality (IR):** it is not harmful for any agent to participate in the market, meaning in any trade there is a strategy that yields a utility that is no less than their initial utility. Note that this says nothing about the outcome of the event – agents may still experience a net loss for having purchased a bet which turned out to be false.
- **Budget Balance (BB):** the sum of all payments is at least zero, meaning no extra funds have to be supplied to subsidise the market.

Now we have all we need to discuss recent results in the literature on algorithmic mechanism design for combinatorial auctions.

3.2 Related academic work

Algorithmic mechanism design enjoys much attention in the setting of combinatorial auctions, where there is much opportunity to make use of traditional techniques from theoretical computer science. Much of this is inspired by a question first posed by Nisan and Ronen [?], which asks whether the requirement for dominant strategy incentive compatibility inherently degrades a mechanism's approximation ratio – put another way, if mechanism design is inherently harder than algorithm design. This question is covered in part by Daniely, Schapira, and Shahaf [4], who present various inapproximability results for deterministic, truthful mechanisms. Specifically, they show that for general valuation functions no computationally-efficient mechanism may achieve better performance than the trivial m-approximation, which collects all items into a single bundle and awards it to the highest bidder for the price of the second-highest bid, while remaining truthful. Colini-Baldeschi et al. [3] make use of randomisation to present three constant-factor approximation mechanisms for two-sided combinatorial markets, and experiment with various assumptions on the structure of the agents' valuations. In particular, they provide three 6-approximation mechanisms which achieve individual rationality, incentive compatibility, and a stronger notion of budget balance which they term Direct Trade Strong Budget Balance (DSBB).

3.3 Existing systems

4 Research Processes

5 Project Management

6 Progress

7 Conclusion

References

[1] *Predictalot! (and we mean a lot).* `http://blog.oddhead.com/2010/03/05/predictalot/`. Accessed: 2020-02-16.

[2] E. H. Clarke, *Multipart pricing of public goods*, Public Choice, 11 (1971), pp. 17–33.

[3] R. Colini-Baldeschi, P. W. Goldberg, B. de Keijzer, S. Leonardi, T. Roughgarden, and S. Turchetta, *Approximately Efficient Two-Sided Combinatorial Auctions*, in Proceedings of the 2017 ACM Conference on Economics and Computation - EC '17, Cambridge, Massachusetts, USA, 2017, ACM Press, pp. 591–608.

[4] A. Daniely, M. Schapira, and G. Shahaf, *Inapproximability of Truthful Mechanisms via Generalizations of the Vapnik–Chervonenkis Dimension*, SIAM Journal on Computing, 47 (2018), pp. 96–120.

[5] T. Groves, *Incentives in Teams*, Econometrica, 41 (1973), p. 617.

[6] N. Nisan and A. Ronen, *Algorithmic Mechanism Design (Extended Abstract)*, in Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, STOC '99, New York, NY, USA, 1999, Association for Computing Machinery, p. 129–140.

[7] ———, *Computationally Feasible VCG Mechanisms*, J. Artif. Int. Res., 29 (2007), p. 19–47.

[8] W. Vickrey, *Counterspeculation, Auctions, and Competitive Sealed Tenders*, The Journal of Finance, 16 (1961), pp. 8–37.