# Table of Contents

Editor’s Introduction 1
SHADDIN DUGHMI

SIGecom News 2
DAVID C. PARKES

Approximately Optimal Mechanism Design: Motivation, Examples, and Lessons Learned 4
TIM ROUGHGARDEN

On Risk Measures, Market Making, and Exponential Families 21
JACOB D. ABERNETHY, RAFAEL M. FRONGILLO, and SINDHU KUTTY

Computational Aspects of Random Serial Dictatorship 26
HARIS AZIZ, FELIX BRANDT, MARKUS BRILL, and JULIÁN MESTRE

A Simple and Approximately Optimal Mechanism for an Additive Buyer 31
MOSHE BABAIOFF, NICOLE IMMORLICA, BRENDAN LUCIER, and S. MATTHEW WEINBERG

Network Improvement for Equilibrium Routing 36
UMANG BHASKAR and KATRINA LIGETT

Spliddit: Unleashing Fair Division Algorithms 41
JONATHAN GOLDMAN and ARIEL D. PROCACCIA

The Cost of Annoying Ads 47
DANIEL G. GOLDSTEIN, R. PRESTON MCAFEE, and SIDDHARTH SURI
Editor’s Introduction

SHADDIN DUGHMI
University of Southern California

This summer I officially took over from Ariel Procaccia as the editor of SIGecom Exchanges. I plan to continue in the tradition of my predecessors; namely, curating a selection of the best research in our community in the form of easily-digestible short articles as well as longer surveys and position papers. As work at the interface of computer science and economics grows both in scope and in sheer quantity, Exchanges can be a low-cost way to keep us all up to speed on the latest and greatest.

Issue 13.2 includes an update from David Parkes on the status of the SIG, a survey on approximately optimal mechanism design by Tim Roughgarden, and six research letters. My goal in soliciting contributions was to represent a variety of the influential threads of research in our community. For future issues, please consider volunteering a letter or survey, or suggesting an area of work you would like to see represented. Additionally, if you have any ideas on how to expand the scope of exchanges so as to promote more engagement with the community, please drop me an email.

Finally, I would like to thank ACM and the SIG for their support, with special thanks to Ariel Procaccia and Felix Fischer for their help with the transition.

Author’s address: shaddin@usc.edu
We have had another successful year! As always, let me ask that you remember to join the SIG\(^1\) to keep our membership healthy and be able to participate in the nominating process for our awards, and upcoming elections for new officers. Also remember to use the ACM Author-Izer to drive additional revenue to the SIG.\(^2\)

I’m happy to report that the latest edition of the ACM Conference on Economics and Computation (newly renamed) was a great success. EC’14 was held at Stanford University, Palo Alto CA in June 2014, and was collocated with a meeting of the NBER Market Design working group and the NSF/CEME Decentralization Conference. The conference again broke all previous attendance and submission records, with more than 220 attendees and 290 submissions. 80 papers were accepted, for an acceptance rate of just under 28%. Congratulations to everyone involved, and special thanks to Moshe Babaioff, Vincent Conitzer and David Easley.

We are now looking forward to the 16th ACM Conference on Economics and Computation, which will take place June 15-19, 2015, at the ACM’s Federated Computing Research Conference (FCRC) in Portland, OR. The conference will partially overlap with the 47th ACM Symposium on Theory of Computing (STOC 2015) and we anticipate some opportunities for collaboration between the conferences. The General Chair for EC’15 is Tim Roughgarden (Stanford University) and the Program Co-Chairs are Michal Feldman (Tel Aviv University) and Michael Schwarz (Google Strategic Technology). The deadline for paper submissions is February 10, 2015. Please consider submitting a workshop or tutorial proposal to EC’15. Workshops are a good way to bring together an interdisciplinary audience around a new set of problems. I’m sure that Arpita Ghosh (workshop chair, ec15-workshops-chair@acm.org) and Liad Blumrosen (tutorial chair, ec15-tutorial-chair@acm.org) would be glad to provide informal feedback about ideas in advance of the February 17, 2015, proposal deadline.

Looking further ahead, initial planning is underway to collocate EC’16 with the 2016 Game Theory World Congress (July 24-28, 2016) in Maastricht, The Netherlands.

Happily, I can announce that we have passed a significant landmark and made

\(^{1}\)http://www.sigecom.org/membership.html

\(^{2}\)By generating an Author-Izer link and including on your website you can allow anyone to download your articles from the ACM Digital Library at no charge, and in doing so provide additional revenue share from the ACM to the SIG.

http://www.acm.org/publications/acm-author-izer-service

Author’s address: parkes@eecs.harvard.edu
the first award this past year of the SIGecom Doctoral Dissertation Award. This year this went to Balasubramanian Sivan (University of Wisconsin) for his thesis, “Prior Robust Optimization.” Runner-up awards went to Yang Cai (MIT) and Sigal Oren (Cornell University). Congratulations to all! We are again soliciting nominations for our two SIG awards:

—The **SIGecom Doctoral Dissertation Award**, which recognizes an outstanding dissertation in the fields of electronic commerce, and economics and computation; and,

—The **SIGecom Test of Time Award**, which recognizes an influential paper or series of papers by a single author or team of authors that has had significant impact on research or applications in the fields of electronic commerce or economics and computation.

**Deadlines for these two awards are February 28, 2015 and March 13, 2015, respectively, and you can see the complete instructions here** [http://www.sigecom.org/awards.html](http://www.sigecom.org/awards.html). Thanks to the two committees that have agreed to adjudicate these awards, and especially to Michael Wellman and David Pennock for agreeing to chair this year’s committees.

This year we continued to support as an in-cooperation event the *Conference on Web and Internet Economics* (WINE), which will take place in Beijing, December 14-17, 2014. The *7th International Symposium on Algorithmic Game Theory* (SAGT), which took place in Patras, Greece, September 30-October 2, 2014, is another in-cooperation event.

In closing, thanks to all of the volunteers that make the activities of the SIG possible. A special shout-out this year to Ariel Procaccia for his wonderful efforts in regard to editing our newsletter, SIG Exchanges, and to Shaddin Dughmi for taking over as editor.

Don’t hesitate to contact me with any questions or suggestions about the SIG, and see you in Portland in 2015!

Yours,
David C. Parkes
President, ACM SIGecom
Approximately Optimal Mechanism Design: Motivation, Examples, and Lessons Learned

TIM ROUGHGARDEN
Stanford University

This survey describes the approximately optimal mechanism design paradigm and uses it to investigate two basic questions in auction theory. First, when is complexity — in the sense of detailed distributional knowledge — an essential feature of revenue-maximizing single-item auctions? Second, do combinatorial auctions require high-dimensional bid spaces to achieve good social welfare?

Categories and Subject Descriptors: F.0 [Theory of Computation]: General
General Terms: Algorithms, Economics, Theory
Additional Key Words and Phrases: Mechanism design, auctions, approximation

1. INTRODUCTION
1.1 Preamble
Optimal mechanism design enjoys a beautiful and well-developed theory, and also a number of killer applications. Rules of thumb produced by the field influence everything from how governments sell wireless spectrum licenses to how the major search engines auction off online advertising.

There are, however, some basic problems for which the traditional optimal mechanism design approach is ill-suited — either because it makes overly strong assumptions, or because it advocates overly complex designs. The thesis of this survey is that approximately optimal mechanisms allow us to reason about fundamental questions that seem out of reach of the traditional theory.

1.2 Organization
This survey has three main parts. The first part reviews a couple of the greatest hits of optimal mechanism design, the single-item auctions of Vickrey and Myerson. We’ll see how taking baby steps beyond these canonical settings already highlights limitations of the traditional optimal mechanism design paradigm, and motivates a more relaxed approach. This part also describes the approximately optimal mechanism design paradigm — how it works, and what we aim to learn by applying it.

The second and third parts of the survey cover two case studies, where we instantiate the general design paradigm to investigate two basic questions. In the

This survey is based on a talk given by the author at the 15th ACM Conference on Economics and Computation (EC), June 2014.
This work is supported in part by NSF grants CCF-1016885 and CCF-1215965, an ONR PECASE Award.
Author’s address: T. Roughgarden, Computer Science Department, Department of Computer Science, 474 Gates Building, Stanford, CA 94705 USA. Email: tim@cs.stanford.edu.
first example, we consider revenue maximization in a single-item auction with heterogeneous bidders. Our goal is to understand if complexity — in the sense of detailed distributional knowledge — is an essential feature of good auctions for this problem, or alternatively if there are simpler auctions that are near-optimal. The second example considers welfare maximization with multiple items. Our goal here is similar in spirit: when is complexity — in the form of high-dimensional bid spaces — an essential feature of every auction that guarantees reasonable welfare? Are there interesting cases where low-dimensional bid spaces suffice?

2. THE OPTIMAL AND APPROXIMATELY OPTIMAL MECHANISM DESIGN PARADIGMS: VICKREY, MYERSON, AND BEYOND

2.1 Example: The Vickrey Auction

Let’s briefly recall the Vickrey or second-price single-item auction [Vickrey 1961]. Consider a single seller with a single item; assume for simplicity that the seller has no value for the item. There are $n$ bidders, and each bidder $i$ has a valuation $v_i$ that is unknown to the seller. Vickrey’s auction is designed to maximize the welfare, which in a single-item auction just means awarding the item to the bidder with the highest valuation. This sealed-bid auction collects a bid from each bidder, awards the item to the highest bidder, and charges the second-highest price. The point of the pricing rule is to ensure that truthful bidding is a dominant strategy for every bidder. Provided every bidder follows its dominant strategy, the auction maximizes welfare ex post (that is, for every valuation profile).

In addition to being theoretically optimal, the Vickrey auction has a simple and appealing format. Plenty of real-world examples resemble the Vickrey auction. In light of this confluence of theory and practice, what else could we ask for? To foreshadow what lies ahead, we mention that when selling multiple non-identical items, the generalization of the Vickrey auction is much more complex.

2.2 Example: Myerson’s Auction

What if we want to maximize the seller’s revenue rather than the social welfare? Since there is no single auction that maximizes revenue ex post, the standard approach here is to maximize the expected revenue with respect to a prior distribution over bidders’ valuations. So, assume bidder $i$’s valuation is drawn independently from a distribution $F_i$ that is known to the seller. For the moment, assume also that bidders are homogeneous, meaning that their valuations are drawn i.i.d. from a known distribution $F$.

Myerson [1981] identified the optimal auction in this context, and under mild conditions on $F$ it is a simple twist on the Vickrey auction — a second-price auction with a reserve price $r$. Moreover, the optimal reserve price is simple and intuitive — it is just the monopoly price $\arg\max_p[p \cdot (1 - F(p))]$ for the distribution $F$, the optimal take-it-or-leave-it offer to a single bidder with valuation drawn from $F$. Thus, to implement the optimal auction, you don’t need to know much about the valuation distribution $F$ — just a single statistic, its monopoly price.

\footnote{That is, the winner is the highest bidder with bid at least $r$, if any. If there is a winner, it pays either the reserve price or the second-highest bid, whichever is larger.}
Once again, in addition to being theoretically optimal, Myerson’s auction is simple and appealing. It is more or less equivalent to an eBay auction, where the reserve price is implemented using an opening bid. Given this success, why do we need to enrich the traditional optimal mechanism design paradigm? As we’ll see, when bidders’ valuations are not i.i.d., the theoretically optimal auction is much more complex and no longer resembles the auction formats that are common in practice.

2.3 The Optimal Mechanism Design Paradigm

Having reviewed two famous examples, let’s zoom out and be more precise about the optimal mechanism design paradigm. The first step is to identify the design space of possible mechanisms, such as the set of all sealed-bid auctions. The second step is to specify some desired properties. In this talk, we focus only on cases where the goal is to optimize some objective function that has cardinal meaning, and for which relative approximation makes sense. We have in mind objectives such as the seller’s revenue (in expectation with respect to a prior) or social welfare (ex post) in a transferable utility setting. The goal of the analyst is then to identify one or all points in the design space that possess the desired properties — for example, to characterize the mechanism that maximizes the welfare or expected revenue.

What can we hope to learn by applying this framework? The traditional answer is that by solving for the optimal mechanism, we hope to receive some guidance about how to solve the problem. With the Vickrey and Myerson auctions, we can take the theory quite literally and simply implement the mechanism advocated by the theory. More broadly, one looks for features present in the theoretically optimal mechanism that seem broadly useful — for example, Myerson’s auction suggests that combining welfare maximization with suitable reserve prices is a potent approach to revenue-maximization.

There is a second, non-traditional answer that we exploit explicitly when we extend the paradigm to accommodate approximation. Even when the theoretically optimal mechanism is not directly useful to the practitioner, for example because it is too complex, it is directly useful to the analyst. The reason is that the performance of the optimal mechanism can serve as a benchmark, a yardstick against which we measure the performance of other designs that are more plausible to implement.

2.4 The Approximately Optimal Mechanism Design Paradigm

To study approximately optimal mechanisms, we again begin with a design space and an objective function. Often the design space is limited by side constraints such as a “simplicity” constraint. For example, we later consider mechanisms with limited distributional knowledge, and those with low-dimensional bid spaces.

The new ingredient of the paradigm is a benchmark. This is a target objective function value that we would be ecstatic to achieve. Generally, the working hypothesis will be that no mechanism in the design space realizes the full value of the benchmark, so the goal is to get as close to it as possible. In the two examples we discuss, where the design space is limited by a simplicity constraint, a simple and natural benchmark is the performance achieved by an unconstrained, arbitrarily complex mechanism. The goal of the analyst is to identify a mechanism in the
design space that approximates the benchmark as closely as possible. For example, it is clearly interesting to establish that there is a “simple” mechanism with performance almost as good as an arbitrarily complex one.

What is the point of applying this design paradigm? The first goal is exactly the same as with the traditional optimal mechanism design paradigm. Whenever you have a principled way of selecting out one mechanism from many, you can hope that the distinguished mechanism is literally useful or highlights features that are essential to good designs. The approximation paradigm provides a novel way to identify candidate mechanisms.

There is a second reason to use the approximately optimal mechanism design paradigm, which has no analog in the traditional approach. The approximation framework enables the analyst to quantify the cost of imposing side constraints on a mechanism design space. For example, if there is a simple mechanism with performance close to that of the best arbitrarily complex mechanism, then this fact suggests that simple solutions might be good enough. Conversely, if every point in the design space is far from the benchmark, then this provides a forceful argument that complexity is an essential feature of every reasonable solution to the problem.

2.5 Two Case Studies

Sections 3 and 4 instantiate the approximately optimal mechanism design paradigm to study two fundamental questions. We first study expected revenue-maximization in single-item auctions, with bidders that have independent but not necessarily identically distributed valuations. The theoretically optimal mechanism can be complex, in the sense that it requires detailed distributional knowledge. We use the approximation paradigm to identify when such complexity is an inevitable property of every near-optimal auction.

Our second case study concerns welfare maximization. Here, the complexity stems from selling multiple non-identical items. Again, the theoretically optimal mechanism is well known but suffers from several drawbacks that preclude direct use. We apply the approximation paradigm to identify when simpler mechanisms, meaning mechanisms with low-dimensional bid spaces, can perform well, versus when complex bid spaces are necessary for non-trivial welfare guarantees.

2.6 Other Applications of the Approximation Paradigm

An enormous amount of research over the past fifteen years, largely but not entirely in the computer science literature, can be viewed as instantiations of the approximately optimal mechanism design paradigm. This survey merely singles out two recent examples that are near and dear to the author’s heart.

For example, all of the following questions have been studied through the lens of approximately optimal mechanisms.

(1) What is the cost of imposing bounded communication in settings with very large type spaces, such as combinatorial auctions? This line of research originated in Nisan and Segal [2006] and is surveyed by Segal [2006].

(2) What is the cost of imposing bounded computation in settings that involve computationally difficult optimization problems, such as combinatorial auctions? Two early papers are Lehmann et al. [2002] and Nisan and Ronen [2001]
and a recent survey is Nisan [2014].

(3) What is the cost of limiting the distributional knowledge of a mechanism? Several papers in the economics literature shed light on this question [Baliga and Vohra 2003; Bulow and Klemperer 1996; Neeman 2003; Segal 2003]. The approximation interpretation is explicit in Dhangwatnotai et al. [2010]; see also Hartline [2013] for a survey. The case study in Section 3 is another example of work in this vein.

(4) Are there auctions that achieve good revenue in the worst case (i.e., ex post)? This question was formalized using the approximately optimal mechanism design framework in Goldberg et al. [2006]; see Hartline [2013] for a recent survey.

(5) Are there mechanisms with “simple” allocation rules that perform almost as well as arbitrarily complex mechanisms? For example, see Chawla et al. [2007] and Hartline and Roughgarden [2009] for revenue guarantees for auctions that make use only of welfare-maximization supplemented by reserve prices.

(6) Are there mechanisms with “simple” pricing rules that perform almost as well as arbitrarily complex mechanisms? See Lucier and Borodin [2010] and Caragiannis et al. [2012] for case studies in combinatorial and keyword auctions, respectively. The case study in Section 4 is another example of this and the preceding directions.

3. CASE STUDY: DO GOOD SINGLE-ITEM AUCTIONS REQUIRE DETAILED DISTRIBUTIONAL KNOWLEDGE?

This section applies the approximately optimal mechanism design paradigm to the problem of revenue-maximization in single-item auctions. The take-away from this exercise is that the amount of distributional knowledge required for near-optimal revenue is governed by the degree of bidder heterogeneity.

3.1 Optimal Single-Item Auctions

We now return to expected revenue-maximization in single-item auctions, but allow heterogeneous bidders, meaning that each bidder i’s private valuation vi is drawn independently from a distribution F_i that is known to the seller. Myerson [1981] characterized the optimal auction, as a function of the distributions F_1, . . . , F_n.

The trickiest step of Myerson’s optimal auction is the first one, where each bid b_i is transformed into a virtual bid φ_i(b_i), defined by

\[ φ_i(b_i) = b_i - \frac{1 - F_i(b_i)}{f_i(b_i)}. \]  

The exact functional form in (1) is not important for this paper, except to notice that computing φ_i(b_i) requires knowledge of the distribution, namely of f_i(b_i) and F_i(b_i).

Given this transformation, the rest of the auction is straightforward. The winner is the bidder with the highest positive virtual bid (if any). To make truthful bidding a dominant strategy, the winner is charged the minimum bid at which it would continue to be the winner.2

2We have only described the optimal auction in the special case where each distribution F_i is reg-
When all the distributions \( F_i \) are equal to a common \( F \), and hence all virtual valuation functions \( \varphi_i \) are identical, the optimal auction simplifies and is merely a second-price auction with a reserve price of \( \varphi^{-1}(0) \), which turns out to be the monopoly price for \( F \). In this special case, the optimal auction requires only modest distributional knowledge — a single statistic, the monopoly price. In general, the optimal auction does not simplify further than the description above, and detailed distributional knowledge is required to compute and compare the virtual bids of bidders with different valuation distributions.

3.2 Motivating Question
This section uses the approximately optimal mechanism design paradigm to study the following question.

*Does a near-optimal single-item auction require detailed distributional knowledge?*

To study this question formally, we need to parameterize the “amount of knowledge” that the seller has about the valuation distributions. We look to computational learning theory, a well-developed branch of computer science [Valiant 1984], for inspiration. We consider a seller that does not know the valuation distributions \( F_1, \ldots, F_n \), except inasmuch as it knows \( s \) valuation profiles \( v^{(1)}, \ldots, v^{(s)} \) that have been sampled i.i.d. from these distributions. In an auction context, an obvious interpretation of these samples is as the valuations of comparable bidders in past auctions for comparable items, as inferred from bid data. See Ostromsky and Schwarz [2009] for a real-world example of this approach, in the context of setting reserve prices in Yahoo! keyword auctions.

Thus, our design space is the set of auctions that depend on the valuation distributions only through samples. Formally, for a parameter \( s \geq 1 \), a point in the design space is a function from \( s \) valuation profiles (the samples) to a single-item auction, which is then run tomorrow on a fresh valuation profile drawn from \( F_1 \times \cdots \times F_n \).

Our objective function is the expected revenue, where the expectation is over both the samples (which determines the auction used) and the final valuation profile (which determines the revenue earned by the chosen auction).

Our benchmark — the highest expected revenue we could conceivably obtain — is simply the expected revenue earned by Myerson’s optimal auction for the distributions \( F_1, \ldots, F_n \). We call this the Myerson benchmark. Thus, we are comparing the optimal expected revenue obtainable by a seller with partial distributional knowledge to that by a seller with full distributional knowledge. The goal is to understand the amount of knowledge (i.e., the number of samples) needed to earn expected revenue at least \((1 - \epsilon)\) times the Myerson benchmark, where \( \epsilon \) is a parameter such as 0.1 or 0.01.

3.3 Formalism: One Bidder
To make sure that the formalism is clear, let’s warm up with a simple example. In addition to only one seller with one item, suppose there is also only one bidder, with

\( \text{universal, meaning that the virtual valuation functions } \varphi_i \text{ are nondecreasing. The general case “monotonizes” the virtual valuation functions — monotonicity is essential for incentive-compatibility — but otherwise applies the same three steps [Myerson 1981].} \)
valued from a distribution $F$ unknown to the seller. With only one bidder, auctions are merely take-it-or-leave-it offers.\(^3\) The goal is to design a function $p(v_1, \ldots, v_s)$ from samples $v_1, \ldots, v_s \sim F$ to prices that, for every $F$, achieves expected revenue $\mathbb{E}_{v_1, \ldots, v_s}[p(v_1, \ldots, v_s) \cdot (1 - F(p(v_1, \ldots, v_s)))$ close to that achieved by the monopoly price $\arg\max_p p \cdot (1 - F(p))$ of $F$. In other words, given data from $s$ past transactions, the goal is to set a near-optimal price for a new bidder encountered tomorrow. See also Figure 1.

### 3.4 Results for a Single Bidder

We next state a series of results for the single-bidder special case. These results are not the main point of this case study, and instead serve to calibrate our expectations for what might be possible for single-item auctions with multiple bidders.

The bad news is that, without any assumptions about the unknown distribution $F$, no finite number of samples yields a non-trivial expected revenue guarantee for every $F$. That is, for every finite $s$, there is a valuation distribution $F$ such that you learn essentially nothing about $F$ from $s$ samples.\(^4\) This observation motivates restrictions on the unknown distribution.

The good news is that under a standard “regularity” condition, intuitively stating that the tail of $F$ is no heavier than a power-law distribution, is sufficient for interesting positive results.\(^5\) Even just one sample can be used to obtain a non-trivial revenue guarantee for unknown regular distributions: for every such $F$, the function $p(v_1) = v_1$ — using yesterday’s bid as tomorrow’s price — yields expected revenue at least 50\% times that of the monopoly price.\(^6\)

---

\(^3\)Probability distributions over take-it-or-leave-it-offers are also allowed. We discuss only deterministic auctions for simplicity of presentation, but the results of this section also apply to randomized auctions.

\(^4\)For example, for a parameter $M \to \infty$, consider distributions $F$ that put a point mass of $1/M$ at $M$ and are otherwise zero.

\(^5\)Formally, a distribution $F$ is regular if its virtual valuation distribution (1) is nondecreasing.

\(^6\)This is a consequence of the following special case of the Bulow-Klemperer theorem on auctions vs. negotiations [Bulow and Klemperer 1996]: the expected revenue of a Vickrey auction with two bidders with valuations drawn i.i.d. from a regular distribution $F$ is at least that of an optimal
What if we want a better revenue guarantee, like 90% or 99% of this benchmark? To achieve a \((1 - \epsilon)\)-approximation guarantee, we expect the number of samples required to increase with \(1/\epsilon\). Happily, the amount of data required is relatively modest, scaling as a polynomial function of \(1/\epsilon\). For an unknown regular distribution, this function is roughly \(\epsilon^{-3}\) [Dhangwatnotai et al. 2010; Huang et al. 2014]. The sample complexity improves if we impose stronger conditions on the tails of the valuation distribution. For example, if \(F\) satisfies the monotone hazard rate condition — meaning \(f(x)/(1 - F(x))\) is nondecreasing in \(x\) — then roughly \(\epsilon^{-3/2}\) samples are necessary and sufficient to achieve a \((1 - \epsilon)\)-approximation of the benchmark [Huang et al. 2014]. The upper bounds on sample complexity follow from natural pricing strategies, such as choosing the monopoly price for the empirical distribution of the samples.

3.5 Formalism: Multiple Bidders

Generalizing the formalism to single-item auctions with multiple bidders proceeds as one would expect. The seller is now given \(s\) samples \(v_1, \ldots, v_s\), where each sample \(v_j\) is a valuation profile, comprising one valuation (drawn from \(F_i\)) for each bidder \(i\). The seller picks an auction \(A(v_1, \ldots, v_s)\) that is a function of these samples only. Recall that the Myerson benchmark is the expected revenue of the optimal auction for \(F_1, \ldots, F_n\). The goal is to design a function \(A(v_1, \ldots, v_s)\) from samples \(v_1, \ldots, v_s \sim F_1 \times \cdots \times F_n\) to single-item auctions that, for every \(F_1, \ldots, F_n\), achieves expected revenue close to this benchmark. As in the single-bidder case, the expectation is over both the past bid data (the samples) and the bidders (a fresh sample from the same distributions). See also Figure 2.

3.6 Positive Results

The hope is that our positive results for the single-bidder problem (Section 3.4) carry over to single-item auctions with multiple bidders. First, provided \(F_1, \ldots, F_n\) are regular distributions, it is still possible to get a coarse but non-trivial approximation (namely, 25%) with a single sample — this follows from a generalization of the Bulow-Klemperer theorem given by Hartline and Roughgarden [2009]. But...
what about very close approximations, like 90% or 99%?

In the special case where bidders are homogeneous — meaning have identically distributed valuations — the positive results for a single bidder continue to hold. Intuitively, the reason is that the form of the optimal auction is independent of the number of bidders — it is simply a second-price auction with a reserve set to the monopoly price for the distribution $F$. Since a single statistic about the distribution $F$ determines the optimal auction for an arbitrary number of homogeneous bidders, it makes sense that the sample complexity of approximating this optimal auction is independent of $n$.

Thus, in these cases, the amount of data — the granularity of knowledge about the valuation distributions — necessary to achieve near-optimal revenue is relatively modest, and does not depend on the number of bidders.

### 3.7 Negative Results

The approximately optimal mechanism design paradigm identifies a qualitative difference between the cases of homogeneous and heterogeneous bidders. When bidders are heterogeneous and we seek a close approximation of the optimal revenue, the sample complexity depends fundamentally on the number of bidders.

**Theorem 3.1 [Cole and Roughgarden 2014].** There is a constant $c > 0$ such that, for every sufficiently small $\epsilon > 0$ and every $n \geq 2$, there is no auction that depends on at most $cn/\sqrt{\epsilon}$ samples and has expected revenue at least $1 - \epsilon$ times the Myerson benchmark for every profile $F_1, \ldots, F_n$ of regular distributions.

The valuation distributions used in the proof of Theorem 3.1 are not pathological — exponential distributions, truncated at different maximum values, already yield the lower bound.

The proof of Theorem 3.1 shows more generally that every auction that fails to implicitly learn all bidders’ virtual valuation functions (recall (1)) up to small error is doomed to having expected revenue less than $1 - \epsilon$ times the Myerson benchmark in some cases. In this sense, detailed knowledge of the valuation distributions is an unavoidable feature of every near-optimal single-item auction with heterogeneous bidders.\(^7\)

### 4. CASE STUDY: DO GOOD COMBINATORIAL AUCTIONS REQUIRE COMPLEX BID SPACES?

In this section we switch gears and study the problem of allocating multiple items to bidders with private valuations to maximize the social welfare. We instantiate the approximately optimal mechanism design paradigm to identify conditions on bidders’ valuations that are necessary and sufficient for the existence of simple combinatorial auctions. The take-away from this section is that rich bidding spaces are an essential feature of every good combinatorial auction when items are complements, while simple auctions can perform well when bidders’ valuations are complement-free.

\(^7\)There is also a converse to Theorem 3.1: for every $\epsilon > 0$ and $n \geq 1$, and for an arbitrary number of bidders with $n$ distinct valuation distributions, a polynomial number (in $n$ and $\epsilon^{-1}$) of samples is sufficient to achieve a $(1 - \epsilon)$-approximation of the Myerson benchmark [Cole and Roughgarden 2014].
4.1 The VCG Mechanism

We adopt the standard setup for allocating multiple items via a combinatorial auction. There are $n$ bidders and $m$ non-identical items. Each bidder has, in principle, a different private valuation $v_i(S)$ for each bundle $S$ of items it might receive. Thus, each bidder has $2^m$ private parameters. In this section, we assume that the objective is to determine an allocation $S_1, \ldots, S_n$ that maximizes the social welfare $\sum_{i=1}^n v_i(S_i)$.

The Vickrey auction can be extended to the case of multiple items; this extension is the Vickrey-Clarke-Groves (VCG) mechanism [Vickrey 1961; Clarke 1971; Groves 1973]. The VCG mechanism is a direct-revelation mechanism, so each bidder $i$ reports a valuation $b_i(S)$ for each bundle of items $S$. The mechanism then computes an allocation that maximizes welfare with respect to the reported valuations. As in the Vickrey auction, suitable payments make truthful revelation a dominant strategy for every bidder.

Even with a small number of items, the VCG mechanism is a non-starter in practice, for a number of reasons [Ausubel and Milgrom 2006]. We focus here on the first step. Every direct revelation mechanism, including the VCG mechanism, solicits $2^m$ numbers from each bidder. This is an exorbitant number: roughly a thousand parameters when $m = 10$, roughly a million when $m = 20$.

4.2 Motivating Question

In this case study, we apply the approximately optimal mechanism design paradigm to study the following question.

*Does a near-optimal combinatorial auction require rich bidding spaces?*

Thus, as in the previous case study, we seek conditions under which “simple auctions” can “perform well.” This time, our design space of “simple auctions” consists of mechanism formats in which the dimension of every player’s bid space is growing polynomially with the number $m$ of items (say $m$ or $m^2$), rather than exponentially with $m$ as in the VCG mechanism.

“Performing well” means, as usual, achieving objective function value (here, social welfare) close to that of a benchmark. We use the *VCG benchmark*, meaning the welfare obtained by the best arbitrarily complex mechanism (the VCG mechanism), which is simply the maximum-possible social welfare.

This case study contributes to the debate about whether or not package bidding is an important feature of combinatorial auctions, a topic over which much blood and ink has been spilled over the past twenty years. We can identify auctions with no or limited packing bidding with low-dimensional mechanisms, and those that support rich package bidding with high-dimensional mechanisms. With this interpretation, our results make precise the intuition that flexible package bidding is crucial when items are complements, but not otherwise.

4.3 A Simple Auction: Selling Items Separately

Our goal is to understand the power and limitations of the entire design space of low-dimensional mechanisms. To make this goal more concrete, we begin by examining a specific simple auction format.
The simplest way of selling multiple items is by selling each separately. Several specific auction formats implement this general idea. We analyze one such format, simultaneous first-price auctions [Bikhchandani 1999]. In this auction, each bidder submits simultaneously one bid per item — only \( m \) bidding parameters, compared with its \( 2^m \) private parameters — and each item is sold in parallel using a first-price auction.

When do we expect simultaneous first-price auctions to have reasonable welfare at equilibrium? Not always. With general bidder valuations, and in particular when items are complements, we might expect severe inefficiency due to the “exposure problem” (e.g., [Milgrom 2004]). For example, consider a bidder in an auction for wireless spectrum licenses that has large value for full coverage of California but no value for partial coverage. When items are sold separately, such a bidder has no vocabulary to articulate its preferences, and runs the risk of obtaining a subset of items for which it has no value, at a significant price.

Even when there are no complementarities amongst the items, we expect inefficiency when items are sold separately (e.g., [Krishna 2010]). The first reason is “demand reduction,” where a bidder pursues fewer items than it truly wants, in order to obtain them at a cheaper price. Second, if bidders’ valuations are drawn independently from different valuation distributions, then even with a single item, Bayes-Nash equilibria are not always fully efficient.

4.4 Valuation Classes

Our discussion so far suggests that simultaneous first-price auctions are unlikely to work well with general valuations, and suffer from some degree of inefficiency even with simple bidder valuations. To parameterize the performance of this auction format, we introduce a hierarchy of bidder valuations (Figure 3); the literature also considers more fine-grained hierarchies [Feldman et al. 2014; Lehmann et al. 2006].

The biggest set corresponds to general valuations, which can encode complementarities among items. The other three sets denote different notions of “complement-free” valuations. In this survey, we focus on the most permissive of these, subadditive.
valuations. Such a valuation $v_i$ is monotone ($v_i(T) \subseteq v_i(S)$ whenever $T \subseteq S$) and satisfies $v_i(S \cup T) \leq v_i(S) + v_i(T)$ for every pair $S, T$ of bundles. This class is significantly larger than the well-studied classes of gross substitutes and submodular valuations.\footnote{Submodularity is the set-theoretic analog of “diminishing returns”: $v_i(S \cup \{j\}) - v_i(S) \leq v_i(T \cup \{j\}) - v_i(T)$ whenever $T \subseteq S$ and $j \notin S$. The gross substitutes condition — which states that a bidder’s demand for an item only increases as the prices of other items rise — is strictly stronger and guarantees the existence of Walrasian equilibria [Kelso, Jr. and Crawford 1982; Gul and Stacchetti 1999].} In particular, subadditive valuations can have “hidden complements” — meaning two items become complementary given that a third item has already been acquired — while submodular valuations cannot [Lehmann et al. 2006].

4.5 When Do Simultaneous First-Price Auctions Work Well?

Our intuition about the performance of simultaneous first-price auctions translates nicely into rigorous statements. First, for general valuations, selling items separately can be a disaster.

**Theorem 4.1** [Hassidim et al. 2011]. With general bidder valuations, simultaneous first-price auctions can have mixed-strategy Nash equilibria with expected welfare arbitrarily smaller than the VCG benchmark.

For example, equilibria of simultaneous first-price auctions need not obtain even 1% of the maximum-possible welfare when there are complementarities between many items.

On the positive side, even for the most permissive notion of complement-free valuations — subadditive valuations — simultaneous first-price auctions suffer only bounded welfare loss.

**Theorem 4.2** [Feldman et al. 2013]. If every bidder’s valuation is drawn independently from a distribution over subadditive valuations, then the expected welfare obtained at every Bayes-Nash equilibrium of simultaneous first-price auctions is at least 50% of the expected VCG benchmark value.

In Theorem 4.2, the valuation distributions of different bidders do not have to be identical, just independent. The guarantee improves to roughly 63% for the special case of submodular bidder valuations [Syrgkanis and Tardos 2013].

Taken together, Theorems 4.1 and 4.2 suggest that simultaneous first-price auctions should work reasonably well if and only if there are no complementarities among items.

4.6 Digression on Approximation Ratios

Before proceeding to our final set of technical results, we pause to emphasize how worst-case approximation results like Theorems 4.1 and 4.2 should be interpreted. Many researchers have a tendency to fixate unduly on and take too literally such approximation guarantees.

Both of the primary motivations for applying the approximately optimal mechanism design paradigm strive for qualitative insights, not fine-grained performance predictions (recall Section 2.4). The first goal is to identify mechanisms or mechanism features that are potentially useful in practice. The auction formats implicitly...
recommended by our case studies, such as selling items separately with first-price auctions provided bidders’ valuations are sufficiently simple, corroborate well with folklore beliefs. The second goal of the approximation paradigm is to quantify the cost of a side constraint like “simplicity” on the mechanism design space. In our case studies, we are coarsely classifying such constraints as “tolerable” or “intolerable” according to whether or not imposing the constraint reduces the achievable performance by a modest constant factor. This viewpoint leads to interesting and sensible conclusions in both of our case studies: complexity is unavoidable in near-optimal revenue-maximizing single-item auctions if and only if bidders are heterogeneous, and complexity is unavoidable in near-optimal welfare-maximizing auctions for selling multiple items if and only if there are complementarities among the items.

To the reader who insists on interpreting approximation guarantees literally, against our advice, we offer a few observations. First, in most applications of the approximately optimal mechanism design framework, the benchmark is constructed so that there is no mechanism in the design space that always achieves 100% of the benchmark. When 100% is unachievable, the best-possible approximation is going to be some number bounded below 100% — it cannot be arbitrarily close to 100% when nothing is tending to infinity. Examples that demonstrate mechanism sub-optimality are often “small” in some sense, which translates to impossibility results for worst-case approximation guarantees better than relatively modest fractions like 50% or, if you’re lucky, 75%. Finally, remember that the benchmark being approximated — for example, the performance of a mechanism so complex as to be unrealizable — is generally not an option on the table. The benchmark represents a utopia that exists only in the analyst’s mind — like your favorite baseball team winning 162 games, or receiving referee reports on your journal submission in less than six months.

Of course, like any general analysis framework, the approximation paradigm can be abused and should be applied with good taste. In settings where the approximately optimal mechanism design paradigm does not give meaningful results, the approach should be modified — by defining a different benchmark, changing the notion of benchmark approximation, or using a completely different analysis framework.

4.7 Negative Results

We now return to the question of when simple mechanisms, meaning mechanisms with low-dimensional bid spaces, can achieve non-trivial welfare guarantees. Section 4.5 considered the special case of simultaneous first-price auctions; here we consider the full design space.

First, the poor performance of simultaneous first-price auctions with general bidder valuations is not an artifact of the specific format: every simple mechanism is vulnerable to arbitrarily large welfare loss when there are complementarities among

\footnote{In some cases it makes sense to speak of the asymptotic optimality of a mechanism, such as the sample complexity results of Section 3 and in large markets (e.g., [Segal 2003; Swinkels 2001]). Asymptotic results are clearly interesting, but are applicable to only a fraction of the problems that we want to reason about.}
items. This impossibility result argues forcefully for a rich bidding language, such as flexible package bidding, in such environments.

**Theorem 4.3 [Roughgarden 2014].** With general bidder valuations, no family of simple mechanisms guarantees equilibrium welfare at least a constant fraction of the VCG benchmark.

In Theorem 4.3, the mechanism family is parameterized by the number of items $m$; “simple” means that the number of dimensions in each bidder’s bid space is bounded above by some polynomial function of $m$. The theorem states that for every such family and constant $c > 0$, for all sufficiently large $m$, there is a valuation profile and a full-information mixed Nash equilibrium of the mechanism with expected welfare less than $c$ times the maximum possible.$^{10}$

We already know from Theorem 4.2 that, in contrast, simple auctions can have non-trivial welfare guarantees with complement-free bidder valuations. Our final result states that no simple mechanism outperforms simultaneous first-price auctions with these bidder valuations.

**Theorem 4.4 [Roughgarden 2014].** With subadditive bidder valuations, no family of simple mechanisms guarantees equilibrium welfare more than 50% of the VCG benchmark.

5. **CONCLUSIONS**

5.1 **Motivating Questions Revisited**

To close the circle, we return to the motivating questions of our case studies and review the answers provided by the approximately optimal mechanism design paradigm. The first question was:

*Does a near-optimal single-item auction require detailed distributional knowledge?*

To answer this question, we took the design space to be auctions with limited knowledge of the valuation distributions — in the form of $s$ i.i.d. samples — and studied the number of samples necessary and sufficient to achieve a $(1 - \epsilon)$-approximation of the Myerson benchmark. We discovered that the amount of knowledge (i.e., samples) required scales linearly with the number of distinct valuation distributions represented in the bidder population. Thus, detailed distributional knowledge is required for near-optimal revenue maximization if and only if the bidders are heterogeneous.

The second motivating question was:

*Does a near-optimal combinatorial auction require rich bidding spaces?*

Here, we defined the design space to be families of mechanisms for which the number of parameters in a bidder’s bid space grows polynomially with the number $m$ of items. We adopted the VCG benchmark, which equals the maximum-possible social welfare. We discovered that high-dimensional bid spaces are fundamental to non-trivial welfare guarantees when there are complementarities among items, but not

---

$^{10}$Technically, Theorem 4.3 proves this statement for an $\epsilon$-approximate Nash equilibrium — meaning every player mixes only over strategies with expected utility within $\epsilon$ of a best response — where $\epsilon > 0$ can be made arbitrarily small. The same comment applies to Theorem 4.4.
otherwise. We also learned that, in some cases, selling items separately with first-price auctions achieves the best-possible worst-case approximation guarantee of any family of simple mechanisms.

5.2 Further Discussion

We showed how the approximately optimal mechanism design paradigm yields basic insights about two fundamental problems. Moreover, it is not clear how to glean these insights without resorting to an analysis framework that incorporates approximation. Our first case study fundamentally involved suboptimality — the less knowledge the seller has, the less revenue it can obtain. Similarly, inefficiency was an unavoidable aspect of our second case study, since simple mechanisms are suboptimal even in very simple settings (e.g., due to demand reduction or bidder asymmetry). An approximation framework is the obvious way to reason about and compare different degrees of suboptimality.

An alternative idea, given a design space and an objective function, is to simply identify the mechanism in the design space with the “best” objective function value. The fundamental issue here is how to meaningfully compare two different mechanisms, which will generally have incomparable performance. For example, for two different single-item auctions that depend on the valuation distributions $F_1, \ldots, F_n$, only through $s$ samples (Section 3), typically either one can have higher expected revenue than the other, depending on the choice of $F_1, \ldots, F_n$. Similarly, for two different combinatorial auctions with low-dimensional bid spaces, one generally has higher welfare for some valuation profiles, and the other for other valuation profiles. The traditional approach in mechanism design to resolving such trade-offs is to impose a prior on the unknown information and maximize expected performance with respect to the prior. But this approach would return us to the very bind we intended to escape, of uninformatively complex optimal mechanisms that require detailed distributional knowledge.

Are our insights surprising? The presented results both confirm some existing intuitions — which we view as important sanity checks for the theory — and go beyond them. For example, in single-item auctions, the result that modest data is sufficient for near-optimal revenue-maximization with homogeneous bidders is natural given that the optimal auction depends only on the valuation distribution’s monopoly price. While revenue-maximization with heterogeneous bidders can only be a more complex problem, it is not clear a priori how such complexity scales with bidder heterogeneity, or even how “complexity” should be defined. The fact that the sample complexity scales linearly with the number of distinct valuation distributions is a satisfying and non-obvious formalization of the idea that “heterogeneity matters.”

For the case study of selling multiple items, the high-level take-aways of our analysis are in line with prevailing intuition — simple auctions enjoy reasonable performance when there are no complementarities among items, but not otherwise. One pleasant surprise of the analysis, which deserves further investigation, is that the positive results for simple auctions hold even for the most general notion of “complement-free valuations,” well beyond the more well-studied special cases of gross substitutes and submodular valuations.
5.3 Open Questions

This survey presented two recent applications of the approximately optimal mechanism design paradigm. There have been dozens of other applications over the past fifteen years (Section 2.6), and there is still much to do.

For example, the sample-complexity formalism of Section 3 shows promise of deepening our understanding of Bayesian-optimal mechanism design. Proving that modest distributional knowledge suffices for near-optimal mechanism performance is an important step in arguing the practical relevance of a theoretically optimal design. Upper bounds on the number of samples needed (Section 3.4) generally suggest interesting methods of incorporating data, such as past bidding data, into designs. Thus far, only the simple settings of single-item auctions (Section 3) and single-bidder multi-item mechanisms [Dughmi et al. 2014] have been studied from this perspective.

For welfare-maximization with multiple items, results like those in Section 4 give preliminary insights into which auction designs might work well, as a function of bidders’ preferences. An important direction for future work is to draw sharper distinctions between different plausibly useful formats. For example, there is ample empirical evidence that ascending auctions for multiple items perform better than their sealed-bid counterparts. Can this observation be made formal using the approximately optimal mechanism design paradigm?

REFERENCES


Huang, Z., Mansour, Y., and Roughgarden, T. 2014. Making the most of few samples. Working paper.
On Risk Measures, Market Making, and Exponential Families

JACOB D. ABERNETHY
University of Michigan
and
RAFAEL M. FRONGILLO
Harvard University
and
SINDHU KUTTY
University of Michigan

In this note we elaborate on an emerging connection between three areas of research: (a) the concept of a risk measure developed within financial mathematics for reasoning about risk attitudes of agents under uncertainty, (b) the design of automated market makers for prediction markets, and (c) the family of probability distributions known as exponential families.

Categories and Subject Descriptors: J.4 [Social and Behavioral Sciences]: Economics; I.2.6 [Artificial Intelligence]: Learning

General Terms: Economics, Theory

Additional Key Words and Phrases: exponential family; entropic risk measure; exponential utility

1. INTRODUCING RISK MEASURES

Imagine that a farmer must decide between several crops to plant for the upcoming growing season. The cost of each crop is different and the yield of each crop depends differently on weather conditions; one may be better suited for cold temperatures, another for heavy rainfall, and yet another for drought. The farmer’s profits for the harvest will thus be determined not only by the cost of the seeds and planting, but also by the suitability of the chosen crops for the actual weather conditions during the season. Given that the weather is uncertain, how should the farmer choose the crops to maximize her profit? Generally speaking, we can model this type of problem by specifying a set $\Omega$ of future states of the world (in this case, the weather during the season), and considering the agent’s position $X : \Omega \to \mathbb{R}$, which specifies the monetary payoff to the agent (in this case, the farmer’s profit) in each such state $\omega \in \Omega$. Now the agent’s actions each induce a different position, and the problem reduces to measuring the quality of such positions and choosing the best.

There are, of course, many ways to evaluate a financial position, including von Neumann-Morgenstern expected utility theory. In this note, we focus on the concept of a (convex) risk measure, which was introduced by the academic finance community [Artzner et al. 1999; Delbaen 2002; Föllmer and Schied 2004] and has appeared more recently in computer science [Othman and Sandholm 2011; Hu and

Authors’ addresses: jabernet@umich.edu, raf@cs.berkeley.edu, skutty@umich.edu
Storkey 2014]. Here, an agent chooses the position which minimizes her risk measure. In general, risk measures are convex functions of positions which satisfy certain axioms, such as monotonicity \((X > X' \Rightarrow \rho(X) \leq \rho(X'))\) and cash invariance \((\rho(X + c\mathbb{1}) = \rho(X) - c\), where \(\mathbb{1}\) denotes the sure payoff of $1\). A major focus of this note will be the entropic risk measure; for any probability measure \(p\) over \(\Omega\), entropic risk is given by:

\[
\rho_p(X) := \log \int_\Omega \exp(-X) dp
\]

The entropic risk measure is related to one popular measure of utility of wealth \(m\), \(u(m) = -\exp(-m)\), commonly known as the exponential utility function. An agent holding some belief distribution \(p\) on \(\Omega\) who maximizes expected exponential utility (under \(p\)) is identical to an agent who minimizes the entropic risk \(\rho_p\). That is, for any two positions \(X, X'\), we have that \(E_p[u(X)] \geq E_p[u(X')] \iff \rho_p(X) \leq \rho_p(X')\). More generally, we can always construct a risk measure from any concave utility function \(u()\) and belief distribution \(p\). As noted by [Föllmer and Schied 2004], we may define a risk measure \(\rho_{u,p}(X) := \inf \{m : E_{\omega \sim p}[u(X(\omega) + m)] \geq u_0\}\). That is, \(\rho_{u,p}(X)\) is the least amount of money the agent needs to maintain expected utility above some default threshold \(u_0\), while holding position \(X\).

It is typical to restrict the space of positions with respect to a payoff function \(\phi: \Omega \rightarrow \mathbb{R}^d\), such that each position \(X\) under consideration can be written \(X(\omega) = r^\top \phi(\omega)\) for some \(r \in \mathbb{R}^d\). Given a fixed \(\phi\), we will often abuse notation and write \(\rho_p(r)\) in place of \(\rho_p(X)\) for the \(X\) defined above, and consider \(r\) to be the “compact” position. As we will see, this compact form lends itself well to the prediction market setting, where the component \(\phi_i(\cdot)\) corresponds to the payout amount for the \(i\)th outcome-contingent contract sold in the market. In addition, for any \(r, q \in \mathbb{R}^d\) it is convenient to define \(\rho_p(r|q) := \rho_p(r + q) - \rho_p(q)\) which may be interpreted as the relative risk of \(r\) given a current position \(q\).

2. MARKET MAKING IN PREDICTION MARKETS

Risk measures provide a surprisingly natural object to design a prediction market via an automated market maker. Prediction markets facilitate aggregation of information via financial incentives, and market designers typically aim to yield accurate predictions of uncertain future events. Goods in these markets correspond to securities with payoffs contingent on some future outcome. The goal is that the prices of these securities should reflect a useful aggregate of information from market participants.

Much attention has been given to the design of automated market makers which facilitate the market by offering to trade with any party at a given price. The task of the market maker is to adjust these prices according to demand. Various formulations of automated market makers have been proposed, such as market scoring rules [Hanson 2003], and the constant-utility market maker [Chen and Pennock 2007]; one that has received considerable attention is the cost function market maker [Abernethy et al. 2013; Chen and Pennock 2007]. In this framework, the market maker possesses a cost function \(C: \mathbb{R}^d \rightarrow \mathbb{R}\) and a current “liability” \(q \in \mathbb{R}^d\); a trader purchasing a bundle of securities \(r \in \mathbb{R}^d\) pays \(C(q + r) - C(q)\) to the market maker, who then updates the liability to \(q + r\).
In [Abernethy et al. 2013], the cost function $C$ is required to satisfy certain axioms (e.g., no arbitrage, expressiveness), which turn out to be essentially the same axioms as those for risk measures mentioned above. Thus, we may equivalently think of the market maker as possessing a compact position $q$ and risk measure $\rho(q) := C(-q)$. Then a trader wishing to purchase bundle $r$ must pay the market maker $\rho(q-r) - \rho(q) = \rho(-r|q)$. It is easy to check that this transaction leaves the market maker’s risk unchanged regardless of $r$, via the cash-invariance principle. Finally, it is interesting to note that the constant-utility market maker of [Chen and Pennock 2007], when viewed as a risk measure, is the same as $\rho_{u,p}$ from above.

3. MARKET SEMANTICS OF EXPONENTIAL FAMILY DISTRIBUTIONS

We will now switch gears to talk about a popular family of probability distributions that turn out to be naturally connected with the entropic risk measure. Given access to empirical averages of some statistics of data, a natural question to ask is if we can find a distribution whose expected statistics match these observations. Exponential family distributions arise as the unique distribution which produces the desired statistics while maximizing Shannon entropy.

Nearly all of the popular probability distributions utilized in the literature can be expressed as an exponential family, including the Gaussian, the multinomial, the Poisson, etc. Let us consider, for example, the Gaussian distribution on a real-valued variable $x$. If instead of the typical parameters of mean $\mu$ and variance $\sigma^2$, we use the natural parameters $\theta = (\mu, \sigma^{-2})$ and we define the vector function $\phi(x) := (x, x^2)$, then we see that the probability density of the Gaussian can be rewritten as $p_\theta(x) \propto \exp(\theta^T \phi(x))$.

The probability density of all exponential families have a similar form, which we now describe. Given a space $\Omega$ and any function $\phi : \Omega \to \mathbb{R}^d$, we can define a probability density for every parameter vector $\theta$ in some feasible set $\Theta$ as

$$p_\theta(\omega) := \exp(\theta^T \phi(\omega) - A(\theta)) \quad \text{where} \quad A(\theta) := \log \int_\Omega \exp(\theta^T \phi(\omega))d\nu. \quad (2)$$

The function $A(\cdot)$ is the normalization factor and is commonly known as the log partition function. It may not be lost on the reader that the definition of $A(\cdot)$ is conspicuously similar to the entropic risk defined in (1). Indeed, recent work [Abernethy et al. 2014; Frongillo 2013] has explored an alternative semantic interpretation of the exponential family framework: one can design a market maker by using the log partition function $A(\cdot)$ as a cost function (equivalently, risk measure). That is, we can imagine a market maker selling $d$ reference securities which pay out according to function $\phi(\omega) \in \mathbb{R}^d$ upon observing the outcome $\omega$. The market maker can interpret its position $q$ as a vector of natural parameters $\theta$, so that when traders request to purchase a share bundle $r \in \mathbb{R}^d$, the market maker charges the trader $A(\theta + r) - A(\theta)$ and, for outcome $\omega$, pays the trader $r^T \phi(\omega)$.

This characterization of the exponential family distribution with market semantics gives rise to a number of nice interpretations:

1. Given the market maker’s position $\theta$, the “market prices” of the $d$ securities announced by the market maker are identical to the mean parameters of $p_\theta$.\(^2\)

\(^2\)Note the change of sign, as risk measures deal with gains whereas cost functions deal with losses.

\(^3\)The mean parameters of a distribution $p$ are defined as $\mu_p := \mathbb{E}_{\omega \sim p}[\phi(\omega)]$. 

(2) We can view the market as simply updating its belief according to new information (i.e. trade), or we can alternatively view the market maker as simply maintaining constant risk according to the entropic risk measure.

(3) We can imagine a trader in this market, with some initial belief \( p_\theta \), who trades to minimize the entropic risk measure \( \rho_{p_\theta}(\cdot) \); as mentioned, this is equivalent to the trader maximizing the expected exponential utility. When the trader invests in a set of shares \( r \), this will clearly affect future investment decisions. But the effect on the trader has two potential interpretations: (i) the agent updates the risk measure to \( \rho_{p_\theta}(\cdot|r) \) or (ii) the agent replaces the original belief \( p_\theta \) with an updated belief \( p_{\theta+r} \). Indeed, the risk measure \( \rho_{p_\theta}(\cdot|r) \) is identical to \( \rho_{p_{\theta+r}}(\cdot) \), suggesting that one’s portfolio and one’s belief parameters are interchangeable quantities within this market framework.

(4) We can imagine a trader who knows the true distribution \( p^* \), and that \( p^* = p_\hat{\theta} \) is a member of the exponential family. If the market maker’s position is currently \( \theta \), then the trader has the potential to earn expected profit in the amount of 
\[
D_A(\theta, \hat{\theta}) = \text{KL}(p_{\hat{\theta}}||p_\theta),
\]
the Kullback-Leibler divergence between \( p_{\hat{\theta}} \) and \( p_\theta \).

The trader achieves this by purchasing \( \hat{\theta} - \theta \) shares.

It is worth noting that many of the above properties hold more generally for other distribution families, in particular the class of generalized exponential families.\(^4\)

4. INFORMATION AGGREGATION IN PREDICTION MARKETS

A central thread of research in the prediction market literature seeks to understand how the market aggregates the information of its participants. Results in this vein depend heavily on the equilibrium concept used, as well as how trader behavior is modeled. Many existing results show natural aggregation properties of the market prices, or equivalently, mean parameters [Wolfers and Zitzewitz 2006; Othman and Sandholm 2010; Frongillo et al. 2012]. Here we will present aggregation results which operate in the share space, or equivalently, in the natural parameters.

Consider a market maker with cost function based on the log partition function (2) as described above. Assuming that traders in this market wish to maximize expected utility with respect to their beliefs, we seek to characterize the market equilibrium, which we define to be the final market state after which no trader wishes to continue trading. It was shown by [Abernethy et al. 2014] that, when each trader \( i \) has exponential utility with risk tolerance parameter \( b_i \) and exponential-family belief parameters \( \hat{\theta}_i \), the market equilibrium becomes

\[
\theta_{\text{final}} = \theta_{\text{init}} + \sum_{j=1}^{n} \delta_j = \frac{\theta_{\text{init}} + \sum_{i=1}^{n} b_i \hat{\theta}_i}{1 + \sum_{i=1}^{n} b_i},
\]

where \( \delta_j \) are trader \( i \)’s security purchases and \( \theta_{\text{init}} \) is the initial market state. In other words, the equilibrium state is a risk-tolerance-weighted average of the natural parameters of the traders and the market maker, with the more risk tolerant traders taking on proportionally more of the final position.

\(^3\)The notation \( D_f(x, y) \) refers to the Bregman divergence defined as \( f(x) - f(y) - \nabla f(y) \top (x - y) \).

\(^4\)Introduced by [Gr"unwald and Dawid 2004], these families are maximum-entropy distributions for entropy functions other than Shannon entropy. For details, see [Frongillo 2013, Chap. 4.3].

This result is quite natural, but appears to depend on the synergy between exponential families and exponential utility. Surprisingly, it is shown in [Barrieu and Karoui 2007; Frongillo and Reid 2014] that this result extends to risk-tolerance families of arbitrary risk measures: if the market maker is risk-constant with risk measure $\rho$, and each trader $i$ seeks to minimize risk measure $\rho_i(X) = b_i \rho(X/b_i)$, then the equilibrium state is again the weighted average given by eq. (3).

While we have characterized the equilibrium state of these markets, it remains to understand how to reach it. In particular, does a more dynamic model of trader activity converge to this equilibrium? We consider a very simple dynamic: at each time step, a trader is selected at random, who computes the optimal trade $\delta_t$ given her current position in the market and the current market state. Then as long as every trader has a nonzero probability of being selected in each round, the unique fixed point of this dynamic is again the market equilibrium state (3). Moreover, we can bound the rate of this convergence: the optimality gap at time $t$, as measured by the sum of risks, is $O(1/t)$. These convergence results continue to hold even beyond the risk-tolerance setting, for any choice of risk measures $\rho_i$ for the traders.

REFERENCES


Frongillo, R. 2013. Eliciting private information from selfish agents.


Computational Aspects of Random Serial Dictatorship

HARIS AZIZ  
NICTA and UNSW  
and  
FELIX BRANDT  
Technische Universität München  
and  
MARKUS BRILL  
Duke University  
and  
JULIÁN MESTRE  
University of Sydney

Two fundamental problems in economics are voting and assignment. In both settings, *random serial dictatorship* is a well-established mechanism that satisfies anonymity, *ex post* efficiency, and strategyproofness. We present an overview of recent results on the computational complexity of problems related to random serial dictatorship.

Categories and Subject Descriptors: F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems; I.2.11 [Distributed Artificial Intelligence]: Multiagent Systems; J.4 [Computer Applications]: Social and Behavioral Sciences—Economics

General Terms: Theory, Algorithms, Economics

Additional Key Words and Phrases: Game Theory, Solutions Concepts, Pareto Optimality, Computational Complexity

1. INTRODUCTION

Two fundamental problems in economics are voting and assignment. In the voting setting, *agents* express preferences over *alternatives* and a *social decision scheme* returns a probability distribution over the alternatives based on the agents’ preferences [Gibbard, 1977]. In the assignment setting, *agents* express preferences over objects, usually called *houses* because only one object is assigned to each agent, and a *random assignment rule* returns a random assignment of the houses specifying the probability with which each house is allocated to each agent [Bogomolnaia and Moulin, 2001, Budish et al., 2013]. In both settings, randomization is crucial to achieve minimal fairness requirements such as anonymity and neutrality.

The mechanism known as *random serial dictatorship (RSD)* gives rise to both a desirable social decision scheme [Gibbard, 1977, Aziz et al., 2013b] and a desirable random assignment rule [Bogomolnaia and Moulin, 2001, Crès and Moulin, 2001].

Authors’ addresses: haris.aziz@nicta.com.au, brandtf@in.tum.de, brill@cs.duke.edu, julian.mestre@sydney.edu.au.
In the voting setting, RSD selects a permutation of the agents uniformly at random and then chooses an alternative by sequentially allowing agents in the permutation to refine the set of feasible alternatives to their most preferred of the remaining alternatives. In the assignment setting, RSD selects a permutation of the agents uniformly at random and then lets one agent after another pick his most preferred of the remaining objects. In both settings, RSD is well-known to satisfy anonymity, strategyproofness, and \textit{ex post} efficiency (i.e., it randomizes over Pareto optimal alternatives/allocations). In fact, it has been conjectured to be the only rule that satisfies these properties [see e.g., Lee and Sethuraman, 2011].

This paper surveys recent computational results regarding the probability of choosing an alternative in the context of voting and the probability of an agent getting a house. There are various reasons why computing the actual probabilities (rather than simply executing the mechanism) is important. In the assignment setting, the resulting probabilities of RSD can be viewed as fractional resource allocations such as in scheduling or other applications [see e.g., Svensson, 1994, Abdulkadiroğlu and Sönmez, 1998, Crès and Moulin, 2001]. Similarly, in voting, the probabilities returned by RSD can be interpreted as fractions of time or other resources allotted to the alternatives. Saban and Sethuraman [2013] mentioned identifying the conditions under which the RSD probabilities can be computed in polynomial time as an open problem. In another paper, Mennle and Seuken [2013] propose random hybrid assignment mechanisms that hinge on the RSD probabilities. Finally, the RSD probabilities have also been used in the design of a recent algorithm for cake cutting [Aziz and Ye, 2014].

2. PRELIMINARIES

A voting problem consists of a set $N = \{1, \ldots, n\}$ of agents having preferences over a finite set $A$ of alternatives where $|A| = m$. The preferences of agents over the alternatives are given by a preference profile $\succeq = (\succeq_1, \ldots, \succeq_n)$ where, for each agent $i \in N$, $\succeq_i$ is a complete and transitive preference relation over $A$ where $\succ_i$ denotes the strict part of the relation and $\sim_i$ denoted the indifference part. A preference relation $\succeq_i$ is linear if $a \succ_i b$ or $b \succ_i a$ for all distinct alternatives $a, b \in A$. For convenience, we will represent preference relations as comma-separated lists in which sets denote indifference classes.

The serial dictatorship rule is defined with respect to a permutation $\pi$ over $N$. It starts with the set of all alternatives and then each agent in $\pi$ successively refines the set of alternatives to the set of most preferred alternatives from the remaining set. RSD returns the serial dictatorship outcome with respect to a permutation that is chosen uniformly at random. If the outcome is not unique, we take the uniform probability distribution over the set of selected alternatives to enforce neutrality.

Example 1 (Illustration of RSD in voting). Consider the following preference profile.

$\succeq_{1:} \{a, b, c\}, d \quad \succeq_{2:} \{b, d\}, a, c \quad \succeq_{3:} \{c, a, b, d\}$

1When preferences are linear, Gibbard [1977] has shown that random dictatorship is the only strategyproof and \textit{ex post} efficient social decision scheme.
For permutation 123, the serial dictatorship outcome is \{b\}. The RSD lottery is 
\[^{a : 0, b : 1/2, c : 1/2, d : 0}\].

An assignment problem is a triple \((N, H, \succ)\), where \(N\) is a set of \(n\) agents, \(H\) is a set of \(m\) houses, and \(\succ = (\succ_1, \ldots, \succ_n)\) is a preference profile that contains, for each agent \(i\), a linear preference relation on the set of houses. Every random assignment specifies, for every agent \(i\) and every house \(h\), the probability \(p_{ih}\) that house \(h\) is assigned to agent \(i\). By the Birkhoff-von Neumann theorem, any such assignment can be obtained by a probability distribution over deterministic assignments. RSD chooses a permutation uniformly at random and then lets the agents sequentially take their most preferred object that has not yet been allocated according to the order of the permutation. It is easily observed that the assignment problem is a special case of the voting problem where the set of alternatives \(A\) is the set of all discrete assignments and the preferences of agents over \(A\) are induced by their preferences over \(H\). Although agents have linear preferences over the houses, they are indifferent among different assignments in which they are allocated the same house.

**Example 2 (Illustration of RSD in assignment).** Consider the following preference profile.
\[\succ_1: a, b, c \quad \succ_2: a, b, c \quad \succ_3: b, a, c\]

Agents 1 and 2 are of the same type since they have identical preferences. For each permutation, one can compute the serial dictatorship outcome. For example, the permutation 123 yields the assignment \(\{\{1, a\}, \{2, b\}, \{3, c\}\}\). The RSD random assignment is given by the following table.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/2</td>
<td>1/6</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td>1/2</td>
<td>1/6</td>
<td>1/3</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2/3</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Although the assignment problem is a special case of the voting problem, this does not imply that positive algorithmic results for voting also hold within the assignment domain. The reason is that the translation of an assignment problem to a corresponding voting problem leads to an exponential blowup in the number of alternatives.

3. **RESULTS**

Recently, Aziz et al. [2013a] have shown that computing the RSD probabilities is \#P-complete both in the voting and in the assignment setting. Independently, Saban and Sethuraman [2013] have shown the same result for the assignment setting.

**Theorem 1.** (Aziz et al. [2013a]) In the voting setting, computing the RSD probabilities is \#P-complete.

**Theorem 2.** (Aziz et al. [2013a], Saban and Sethuraman [2013]) In the assignment setting, computing the RSD probabilities is \#P-complete.
As a corollary, checking whether a given RSD probability is greater than or equal to some fixed \( q \in (0,1) \) is NP-hard. The reason is that since an RSD probability takes at most \( n! + 1 \) values, a polynomial-time algorithm for the problem can be used along with binary search to compute the exact probability in polynomial time. In the assignment setting, even checking whether an agent obtains a house with positive probability is intractable.

**Theorem 3.** (Saban and Sethuraman [2013]) In the assignment setting, checking whether an agent gets a house with positive RSD probability is NP-complete.

A corollary of Theorem 3 is that the problem of computing RSD probabilities does not admit an FPRAS (fully polynomial-time randomized approximation scheme) unless the complexity class NP is equal to the complexity class RP [Saban and Sethuraman, 2013].

In contrast to the result in the assignment setting, Aziz et al. [2013a] showed that in the voting setting, the support of the RSD lottery can be computed in polynomial time. The algorithm greedily builds up a permutation, if possible, in which an agent is added to the partial permutation as long as the agent still keeps the target alternative in contention for being selected.

**Theorem 4.** (Aziz et al. [2013a]) In the voting setting, there is a polynomial-time algorithm to check whether an alternative receives positive RSD probability.

The negative complexity results by Saban and Sethuraman [2013] and Aziz et al. [2013a] prompted the need to identify conditions under which RSD probabilities can be computed efficiently. A problem with parameter \( k \) belongs to the class FPT, and is said to be fixed-parameter tractable, if there exists an algorithm that solves every instance \( I \) of the problem in time \( f(k) \cdot \text{poly}(|I|) \), where \( f \) is some computable function independent of \( I \) and \( \text{poly} \) is a polynomial. Aziz and Mestre [2014] examined the parameterized complexity of the problems with respect to the following parameters: the number of agent types, the number of alternatives, the number of alternative types, and the number of houses. Agents have the same type if they have identical preferences. Alternatives have the same type if each agent is indifferent among all of them.

**Theorem 5.** (Aziz and Mestre [2014]) In the voting setting,

1. There is an FPT algorithm for computing the RSD probabilities with parameter \( T = \# \) of agent types.
2. There is an FPT algorithm for computing the RSD probabilities with parameter \( q = \# \) of alternative types.

**Theorem 6.** (Aziz and Mestre [2014]) In the assignment setting,

1. There is an FPT algorithm for computing the RSD probabilities with parameter \( m = \# \) of houses.
2. There is a polynomial-time algorithm for computing the RSD probabilities if the number \( T \) of agent types is bounded.
4. OPEN PROBLEMS

The recent work on computational aspects of random serial dictatorship gives rise to some interesting problems. In the assignment setting, Saban and Sethuraman [2013] proved that computing $RSD$ probabilities does not admit an FPRAS unless $\text{NP} = \text{RP}$. For voting, it remains open whether there exists an FPRAS.

In the assignment setting, Aziz and Mestre [2014] proved that there is a polynomial-time algorithm if the number of agent types is bounded. It remains open whether there exist an FPT algorithm for computing the $RSD$ probabilities for parameter $\#$ of agent types.

Finally, identifying fast exponential-time algorithms for computing the $RSD$ probabilities in both the assignment and the voting setting is another direction for future work.

Acknowledgments. This material is based on work supported by Deutsche Forschungsgemeinschaft under grant BR 2312/7-2 and by a Feodor Lynen research fellowship of the Alexander von Humboldt Foundation. NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

REFERENCES


A Simple and Approximately Optimal Mechanism for an Additive Buyer

MOSHE BABAIOFF
Microsoft Research

and

NICOLE IMMORLICA
Microsoft Research

and

BRENDAN LUCIER
Microsoft Research

and

S. MATTHEW WEINBERG
Princeton University

In this letter we briefly survey our main result from [Babaioff el al. 2014]: a simple and approximately revenue-optimal mechanism for a monopolist who wants to sell a variety of items to a single buyer with an additive valuation.

Categories and Subject Descriptors: J.4 [Social and Behavioral Science]: Economics

General Terms: Economics, Theory

Additional Key Words and Phrases: Optimal Mechanisms, Simple Mechanisms, Revenue, Approximation

1. INTRODUCTION

Imagine that a monopolist seller has a collection of $n$ indivisible items for sale. How should he sell the items to maximize revenue given that the buyers are strategic? With just a single item for sale to a single buyer with value drawn from a distribution $F$, Myerson [1981] shows that the optimal protocol is straightforward: the seller should post a fixed take-it-or-leave-it price $p$ chosen to maximize $p \cdot (1 - F(p))$, the expected revenue (price times probability of sale). Despite the simplicity of the single-item case, extending this solution to handle multiple items remains the primary open challenge in mechanism design.

Consider even the simplest multi-item scenario [Hart and Nisan 2012]: there is a single buyer with item values drawn independently from distributions $D_1, \ldots, D_n$, and whose value for a set of items is additive. Even when there are only two items for sale, it is known that the revenue-optimal mechanism may involve randomiza-

---

1Note that if the seller has unlimited copies of each item for sale, then a mechanism for a single buyer directly extends to the case of multiple buyers.

Authors' addresses: moshe@microsoft.com, nicimm@microsoft.com, brlucier@microsoft.com, smweinberg@csail.mit.edu
tion [Thanassoulis 2004], even to the extent of offering the buyer a choice among infinitely many lotteries [Daskalakis et al. 2014; Hart and Nisan 2013]. This is troubling not only from the perspective of analyzing optimal mechanisms, but also from the point of view of their usefulness. For a mechanism to be useful in practice, it should be simple to describe and transparent in its execution. Indeed, Myerson’s single-item auction is exciting not only for its optimality, but also its practicality. The danger, then, is that revenue-optimal but complex mechanisms for multiple items may share the fate of other mathematically optimal designs, such as the Vickrey-Clarke-Groves mechanism, which are only rarely used in practice [Ausubel and Milgrom 2006]. It is therefore crucial to pair the study of revenue optimization with an exploration of the power of simple auctions. In other words, what is the relative strength of simple versus complex mechanisms?

2. THE FAILURE OF SELLING SEPARATELY OR BY BUNDLING

The above question was posed in general by Hartline and Roughgarden [2009], and by Hart and Nisan [2012] specifically for the setting of a single additive buyer. They proposed the following suggestion for a simple multi-item auction: sell each item separately, posting a fixed price on each one. The optimal price to set on item $i$ would be $\arg\max_p p(1 - D_i(p))$, mirroring the single-item scenario. At first glance, one might expect this simple approach to be optimal. After all, the buyer’s value for each item is sampled independently, and her value for item $i$ doesn’t depend at all on what other items she receives due to additivity. There is absolutely no interaction between the items at all from the buyer’s perspective, so why not sell the items separately? Somewhat counter-intuitively, it turns out that this mechanism might not be optimal. For example, suppose that there are 2 items and the buyer’s value for each item is distributed uniformly on $\{1, 2\}$. Then the optimal price to set on a single item is either 1 or 2, yielding a per-item revenue of 1 and a total revenue of 2. However, the seller could instead post a single take-it-or-leave-it price of 3 on the bundle of both items. Now the consumer is willing to purchase with probability $3/4$, leading to a total revenue of $9/4 > 2$. Hart and Nisan [2012] show that replacing two items whose values are drawn from the uniform distribution with $n$ items whose values are drawn from the Equal-Revenue distribution yields an example with a gap of $\Omega(\log(n))$.

What is going on in this example? The inherent problem is that the buyer’s value for all items concentrates around its expectation. This is potentially helpful for revenue generation, but the strategy of selling items separately cannot exploit this property. On the other hand, the mechanism designed to target such concentration (selling only the grand bundle at a fixed price) does very poorly in settings where concentration doesn’t occur; Hart and Nisan show that this grand-bundle mechanism achieves only an $\Omega(n)$ approximation to the optimal revenue in general. For example, consider an $n$-item instance where the buyer’s value for item $i$ is $2^i$ with probability $1/2^i$, and 0 otherwise. Then one can see that for all $p$, the buyer is only willing to purchase the grand bundle at price $p$ with probability at most $2/p$ and therefore selling the grand bundle yields expected revenue at most 2. Yet, selling each item separately at price $2^i$ yields expected revenue 1 per item and $n$ in $\Omega(\log(n))$. What is going on in this example? The inherent problem is that the buyer’s value for all items concentrates around its expectation. This is potentially helpful for revenue generation, but the strategy of selling items separately cannot exploit this property. On the other hand, the mechanism designed to target such concentration (selling only the grand bundle at a fixed price) does very poorly in settings where concentration doesn’t occur; Hart and Nisan show that this grand-bundle mechanism achieves only an $\Omega(n)$ approximation to the optimal revenue in general. For example, consider an $n$-item instance where the buyer’s value for item $i$ is $2^i$ with probability $1/2^i$, and 0 otherwise. Then one can see that for all $p$, the buyer is only willing to purchase the grand bundle at price $p$ with probability at most $2/p$ and therefore selling the grand bundle yields expected revenue at most 2. Yet, selling each item separately at price $2^i$ yields expected revenue 1 per item and $n$ in $\Omega(\log(n))$.

$^2$The Equal-Revenue distribution has CDF $F(x) = 0$ for $x \leq 1$, and $F(x) = 1 - 1/x$ for $x \geq 1$. 

total. Unfortunately, we must conclude that neither selling separately nor selling together can always approximate the optimal revenue to within a constant factor.

3. OUR RESULT

Our main result is that the maximum of the revenue generated by these two approaches — either selling all items separately or selling only the grand bundle — is a constant-factor approximation to the optimal revenue. In other words, for any product distribution of buyer values, either selling items separately approximates the optimal revenue to within a constant factor, or else bundling all items together does. Since a good approximation to the expected revenue of each approach can be computed in polynomial time given an appropriate access to the distribution, our results furthermore imply a polytime constant-factor approximation mechanism for the case of an additive buyer with independently (and non-identically) distributed values, even without the restriction of simplicity. Moreover, prior to our work, it was not even known if any deterministic mechanism could achieve a constant-factor approximation to the optimal mechanism, even without regard for simplicity or computational efficiency.

Main Result (Informal). In any market with a single additive buyer and arbitrary independent item value distributions, either selling every item separately or selling all items together as a grand bundle generates at least a constant fraction of the optimal revenue.³

This result complements an active research area aimed at characterizing distributions and valuations in which simple mechanisms are precisely optimal [Alaei et al. 2013; Hart and Nisan 2012; Pavlov 2011; Tang and Wang 2014]. In contrast to that literature, we show that a maximum over simple mechanisms is approximately optimal, for arbitrary product distributions and additive valuations. Our result also echoes a similar line of investigation for markets with unit-demand valuations in which a buyer’s value for a set of items is his maximum value for an item in the set. In this setting, it is known [Chawla et al. 2007; Chawla et al. 2010; Chawla et al. 2010] that selling items separately achieves a constant approximation to the optimal revenue. Our result illustrates that a similar approximation can be achieved for additive buyers, provided that we also consider selling all items together as a grand bundle.

To obtain some intuition into our result, recall the example above with \( n \) items and uniformly-distributed values. This example illustrates that selling all items separately may be a poor choice when the value for the grand bundle concentrates around its expectation. What we show is that, in fact, this is the only scenario in which selling all items separately is a poor choice. We prove that if the total value for all items does not concentrate, then selling separately must generate a constant fraction of optimal revenue.

Our argument makes use of a core-tail decomposition technique introduced by Li and Yao [2013] to study the revenue of selling items separately. Roughly speaking, the idea is to split the support of each item’s value distribution at some cutoff into

³The best constant currently known is 6. A lower bound of 2 is also known.
a “tail” (those values that are sufficiently large), and a “core” (the remainder). One can then try to analyze the revenue contributed from items that lie in the tail separately from those that lie in the core. As the cutoffs get larger and larger, the contribution from items in the tail becomes easier to approximate. When the cutoffs are sufficiently large, we can in fact show that selling separately obtains a constant fraction of the tail’s contribution, as it becomes extremely unlikely that multiple items are simultaneously in the tail (and selling separately is optimal for a single item). Similarly, as the cutoffs get smaller and smaller, the contribution from items in the core becomes easier to approximate. When the cutoffs are sufficiently small, we can in fact show that the better of selling separately and together obtains a constant fraction of the core’s contribution, as the contribution from items in the core must concentrate whenever the expected sum of values greatly exceeds the cutoffs. Our result follows by delicately balancing asymmetrically the cutoffs for all items so that both arguments hold simultaneously.

4. CONCLUSIONS AND FUTURE WORK

Our work provides a simple, approximately optimal mechanism for a monopolist seller facing an additive buyer [Babaioff et al. 2014]. Specifically, we show that the better of selling all items separately or all items together achieves a constant-factor approximation. Our analysis shows that this constant factor is at most 6, yet the worst known examples only exhibit a gap of 2. It is important to get a better understanding of where the gap actually lies.

A more pressing follow-up question is extending our results to multiple bidders. Exciting recent work by Yao [2015] provides a lookahead reduction for the case of many additive bidders. Applying our single-buyer mechanisms within his framework, Yao is able to also develop simple constant-factor approximations for auctions with many additive buyers, largely resolving this direction. However, the techniques at hand and in Yao’s reduction appear highly specialized to additive bidders, and generalizing even slightly seems highly non-trivial. Therefore, an intriguing open question is to generalize our results and Yao’s reduction beyond the case of additive bidders.

REFERENCES


Network Improvement for Equilibrium Routing

UMANG BHASKAR
University of Waterloo

and

KATRINA LIGETT
California Institute of Technology

Routing games are frequently used to model the behavior of traffic in large networks, such as road networks. In transportation research, the problem of adding capacity to a road network in a cost-effective manner to minimize the total delay at equilibrium is known as the Network Design Problem, and has received considerable attention. However, prior to our work, little was known about guarantees for polynomial-time algorithms for this problem. We obtain tight approximation guarantees for general and series-parallel networks, and present a number of open questions for future work.

Categories and Subject Descriptors: F.2.0 [General]: Analysis of Algorithms and Problem Complexity—Approximation Algorithms

General Terms: Algorithms; Theory; Economics

Additional Key Words and Phrases: Routing games, network design, Wardrop equilibrium

1. INTRODUCTION

Routing games model network traffic in applications where users choose paths in the network to minimize their delay. The cost, or total delay, at equilibria in routing games is very well-studied: tight bounds are known on the price of anarchy, as well as in many cases on techniques to influence player strategies (e.g., tolls, Stackelberg strategies) to improve the total delay at equilibria. In common practical applications, the primary means of improving the total delay is to add capacity to the network. However, techniques to compute the optimal allocation of additional capacity are primarily heuristic or computationally inefficient. Our work addresses the problem of adding capacity to a network under a budget to improve the total delay at the resulting equilibrium, and obtains guarantees for computationally efficient techniques [Bhaskar et al. 2014].

The problem of figuring out where to add capacity in a network is clearly a non-trivial computational problem: Braess’ paradox is a prominent example where adding capacity can actually worsen the total delay at equilibrium. Given the importance of this problem for the optimization of road networks, it is unsurprising that it has received considerable attention in the literature on transportation research, where a general version of the problem discussed in this article is called the Continuous Network Design Problem [Abdulaal and LeBlanc 1979]. While a number of algorithms for the problem are given in the literature, these algorithms either offer no performance guarantees on the solution obtained, or are highly inefficient computationally (see, e.g., [Yang and H. Bell 1998; Luathep et al. 2011];...
Li et al. 2012]). Prior to our work, little was known about the performance of computationally efficient algorithms.

2. PROBLEM DEFINITION

Formally, a routing game is defined by a directed graph $G = (V, E)$, a vector of non-decreasing delay functions on the edges $(l_e)_{e \in E}$, and a set of triples $\{(s_i, t_i, d_i)\}_{i \in [k]}$ where each triple denotes the source, sink, and demand for a commodity. We assume players are nonatomic: each player controls infinitesimal traffic. Players of commodity $i$ choose a route from $s_i$ to $t_i$, and the (infinite) set of such players constitute the demand for commodity $i$. A strategy profile for the players corresponds to a multicommodity flow $f$, and given a flow $f$, each player of commodity $i$ picks an $s_i$-$t_i$ path $P$ to minimize her delay $\sum_{e \in P} l_e(f_e)$.

A multi-commodity flow $f := (f^i)_{i \in [k]}$ where each $f^i$ is an $s_i$-$t_i$ flow of value $d_i$ is called a valid flow. A valid flow $f$ is called a Wardrop equilibrium if for all $i \in [k]$ and any $s_i$-$t_i$ paths $P$, $Q$ with $f^i_P > 0$, $\sum_{e \in P} l_e(f_e) \leq \sum_{e \in Q} l_e(f_e)$. The Wardrop equilibrium can also be obtained as the solution to a convex program. If all delay functions are strictly increasing, the program is strictly convex, and hence the equilibrium is unique.

To model network improvement, every edge has a capacity $c_e(\beta_e) := c^0_e + \mu_e \beta_e$, where $\beta_e$ is the money allocated to improving the edge (initially zero), $c^0_e$ is the initial capacity of the edge, and $\mu_e$ is the marginal rate of improvement. The latency function on every edge then has the form $l_e(x, \beta_e) = (x/c_e(\beta_e))^{\mu_e} + b_e$. For a given budget $B \in \mathbb{R}_+$, a valid allocation is a vector $\beta = (\beta_e)_{e \in E} \geq 0$ with $\sum_e \beta_e \leq B$.

For a valid allocation $\beta$, let $f(\beta)$ be the resulting equilibrium with delay functions $l_e(x, \beta_e)$. The network improvement problem can then be written as the following mathematical program:

$$
\min_{f, \beta} \sum_e f_e l_e(f_e, \beta_e) \quad \text{s.t.} \quad \beta \geq 0, \quad \sum_e \beta_e \leq B, \quad \text{and} \quad f = f(\beta).
$$

In the program above, both the flow $f$ and the allocation $\beta$ are treated as variables, although the constraints restrict $f$ to be the unique equilibrium flow. Despite being continuous, the objective function in (1) is non-convex. The feasible region may also be disconnected; both of these make the problem difficult to solve. Despite this, our results give tight approximation guarantees for this problem in a number of cases.

3. OUR CONTRIBUTION

We first consider the problem in general graphs. Here, we show that a simple algorithm that operates by relaxing the constraint that $f$ be an equilibrium flow in (1) and solves the resulting convex optimization problem, in fact obtains the best possible approximation ratio for instances with affine delay functions ($n_e = 1$ on all edges).\(^{1}\)

**Theorem 3.1.** We can obtain in polynomial time a $4/3$-approximate allocation for instances with affine delay functions, and an $O(p/\log p)$-approximate allocation

\(^{1}\)The results of Theorem 3.1 were also independently obtained by Gairing, Harks and Klimm [Gairing et al. 2014];

for instances with delay functions of degree $p$. It is NP-hard to obtain an approximation ratio better than $4/3$, even in single-commodity instances with affine delays.

The somewhat surprising optimality of the simple algorithm described is proved by showing connections to previous work on routing games in algorithmic game theory. The proof of the approximation ratio uses well-known bounds on the price of anarchy [Roughgarden 2003]. The proof of hardness uses techniques developed, motivated by Braess’ paradox, for the problem of removing capacity in a network to reduce the total delay at equilibrium [Roughgarden 2006].

Given the practical relevance of the problem, an immediate question is if in special cases, the $4/3$-approximation can be improved upon. In order to circumvent the inapproximability, it is obvious that we must consider restricted topologies. We hence consider the problem in series-parallel graphs. Here, we show that in the single-commodity setting, we can in fact get arbitrarily close to the optimal solution in polynomial time (i.e., get an FPTAS).

**Theorem 3.2.** For single-commodity routing games on series-parallel graphs with polynomial delay functions, for any $\epsilon > 0$, we can obtain a $(1 + \epsilon)$-approximate allocation in time polynomial in the size of the input and $1/\epsilon$.

Our algorithm for Theorem 3.2 uses dynamic programming as well as the structure of series-parallel graphs to obtain the FPTAS. The algorithm consists primarily of three steps. The first step is to modify the mathematical program (1) in two ways. Firstly, we relax the equilibrium constraint in (1) as we did in the proof of Theorem 3.1. Secondly, instead of minimizing the total delay, we modify our objective to minimize the maximum delay over all paths with strictly positive flow. This new objective is a discontinuous function of the flows. However, we show that in series-parallel graphs, the two problems are equivalent: an optimal solution to the second problem is optimal for the original problem as well. The advantage of modifying the problem in this way is to simplify the feasible set. Instead of all valid allocations and flows at equilibrium, our feasible set is now the set of all valid allocations and valid flows.

We then show that we can appropriately discretize this simplified feasible set and search in this discretized space using dynamic programming to obtain a near-optimal solution to the modified problem. Finally, we show that a near-optimal solution to the modified problem is a near-optimal solution to the original problem as well. A number of technical issues arise in the discretization, such as edges where the marginal rate of improvement is large (and thus a small allocation significantly improves the capacity). We show that these can be dealt with, and obtain a polynomial-time algorithm.

A consequence of Theorem 3.2 is that restricting network topologies can circumvent the $4/3$-inapproximability. However, we show that even in series-parallel graphs, the FPTAS is the best possible: the problem is NP-hard. In fact, we show

---

2Series-parallel graphs are defined recursively. A single edge $e = \langle s, t \rangle$ is a series-parallel graph, with terminals $s$ and $t$. Two series-parallel graphs can be combined to obtain a new series-parallel graph, either by merging the $s$-terminals together and the $t$-terminals together (a parallel-join) or by merging the $t$-terminal of one graph with the $s$-terminal of the other (a series-join).
NP-hardness for a special case of series-parallel graphs, obtained by joining in series subgraphs consisting of two parallel edges. The proof of hardness is via reduction from PARTITION.

**THEOREM 3.3.** The network improvement problem is NP-hard, even for single-commodity routing games on series-parallel graphs with affine delay functions.

4. OPEN PROBLEMS

Our work [Bhaskar et al. 2014] presents tight bounds for the settings we consider. There are many open problems and generalizations that follow, and are of interest both theoretically and practically. We present two questions immediately motivated by applications.

**Variable demand.** If the capacity of a network were expanded, it is natural to suppose that more traffic would use the network. This is called “induced demand,” and is frequently a practical concern when expanding capacity in road networks. The price of anarchy with induced demand has been studied previously [Cole et al. 2012]. An open problem is to develop algorithms for network improvement that address induced demand as well.

**Other network topologies.** Our reduction that establishes a lower bound of $4/3$ on the approximation guarantee is not applicable if we restrict the graph topology, e.g., if the graph is planar. However, road networks are often planar or nearly so. Can a better approximation guarantee be obtained for planar graphs?

More generally, there are numerous problems in economics and engineering that are naturally expressed as mathematical programs with equilibrium constraints (see, e.g.,[Luo et al. 1996]); network improvement is just one such problem. Such problems are computational in nature, and have received considerable attention in fields such as operations research and engineering. We believe that, as for the network design problem, the tools of algorithmic game theory can contribute to a better understanding of these natural and well-motivated problems.

**Acknowledgments.** K. Ligett gratefully acknowledges support of an NSF CAREER Award (1254169), the Charles Lee Powell Foundation, and a Microsoft Research Faculty Fellowship.

**REFERENCES**


Spliddit: Unleashing Fair Division Algorithms

JONATHAN GOLDMAN
and
ARIEL D. PROCACCIA
Carnegie Mellon University

Spliddit is a first-of-its-kind fair division website, which offers provably fair solutions for the division of rent, goods, and credit. In this note, we discuss Spliddit’s goals, methods, and implementation.

Categories and Subject Descriptors: J.4.a [Social and Behavioral Sciences]: Economics
General Terms: Algorithms; Design; Human Factors; Economics
Additional Key Words and Phrases: Fair Division

1. OVERVIEW

The origins of the mathematically rigorous study of fair division can be traced back to the work of Hugo Steinhaus during World War II [Steinhaus 1948]. Over the decades, fair division theory has become a major field of study in mathematics, economics, computer science, and political science [Brams and Taylor 1996; Robertson and Webb 1998; Moulin 2003; Procaccia 2013]. Nowadays the literature encompasses provably fair solutions for a wide variety of problems — many of them relevant to society at large. But, to date, very few fair division methods have been made publicly available. Exceptions that prove the rule include the Adjusted Winner Website1, which provides access to a (patented) method for dividing indivisible goods between two players, due to Brams and Taylor [1996]; and Francis Su’s Fair Division Calculator2, which implements methods for splitting rent, divisible goods, and chores from his beautifully written paper [Su 1999]. (Su’s rent division calculator was recently updated by the New York Times3.)

Enter Spliddit (www.spliddit.org), a new fair division website. Quoting from the website:

Spliddit is a not-for-profit academic endeavor. Its mission is twofold:
—To provide easy access to carefully designed fair division methods, thereby making the world a bit fairer.
—To communicate to the public the beauty and value of theoretical research in computer science, mathematics, and economics, from an unusual perspective.

1http://www.nyu.edu/projects/adjustedwinner
2https://www.math.hmc.edu/~su/fairdivision/calc

Authors’ addresses: Computer Science Department, Carnegie Mellon University
{jagoldma,arielpro}@andrew.cmu.edu
Spliddit was officially launched on November 4, 2014. Building on early press coverage in popular technology and science websites such as Gizmodo, Lifehacker, and Slashdot, Spliddit has attracted almost 20,000 visitors in the first week after launch. We expect to serve hundreds of thousands of users in the coming months.

2. THE THREE APPLICATIONS

Spliddit currently includes three applications (see Figure 1), which were selected to maximize broad appeal and usability. A single algorithm is provided for each application, even in cases where several incomparable approaches are present in the fair division literature. This nontrivial design choice is partly driven by usability considerations, and partly by educational considerations: it allows us to focus on giving an accessible explanation (enhanced by animations) of the guaranteed fairness properties in the context of each of the three applications.

In the remainder of this section we describe Spliddit’s applications, as well as their corresponding algorithms and fairness guarantees.

**Splitting rent.** The rent division problem involves $n$ players (housemates) and rooms. The total rent, say $R$, is also given. The goal is to fairly assign the rooms to players and divide the rent.

We assume that each player $i$ has value $V_{ij}$ for room $j$, so that for all $i$, $\sum_j V_{ij} = R$. Suppose that price of room $j$ is $p_j$, and $\sum_j p_j = R$. The utility functions of the players are quasi-linear in prices, that is, the utility of player $i$ for room $j$ is $V_{ij} - p_j$. Therefore, the only information Spliddit needs to elicit from each player is her value $V_{ij}$ for each room $j$ (see Figure 2).

Spliddit’s rent division scheme is an implementation of an algorithm developed by Abdulkadiroğlu et al. [2004]. In a nutshell, it is a market-based algorithm which iteratively increases the prices of overdemanded rooms (at the same rate) and decreases the prices of the other rooms (at the same rate). The allocation is guaranteed to be envy-free: If player $i$ is assigned room $j$ for price $p_j$, then for any room $j'$, $V_{ij} - p_j \geq V_{ij'} - p_{j'}$. In this domain, every envy-free allocation is also Pareto efficient, that is, there is no other allocation for which all players have weakly greater utility, and at least one player has strictly greater utility.

---

---

*Spliddit deals with rooms that are shared by multiple roommates by creating several copies of the room, one for each roommate.*

Interestingly, Brams and Kilgour [2001] observe that there are situations where some prices are negative at every envy-free allocation — some players are paid to live in the house! But the algorithm of Abdulkadiro˘ glu et al. [2004] guarantees that prices will be nonnegative if an envy-free allocation with nonnegative prices exists. In practice, we believe it is highly unlikely that the algorithm will output negative prices for inputs that reflect real-life values.

Dividing goods. Spliddit’s second application allocates a set of goods \( G \) to a set of \( n \) players. The goods can be divisible or indivisible, but the problem is nontrivial only when some of the goods are indivisible. Inheritance is the paradigmatic use case, e.g., dividing an art or jewelry collection (consisting of indivisible goods) among three or more heirs. Each player \( i \) has value \( V_{ij} \) for good \( j \). We assume that valuations are additive, that is, the value of player \( i \) for a bundle of goods \( X \) is \( V_i(X) \triangleq \sum_{j \in X} V_{ij} \). Additivity underlies our fairness guarantees, and, perhaps even more importantly, it facilitates easy elicitation of preferences. On Spliddit, each player simply distributes a pool of 1000 points between the goods.

Let us denote an allocation of the goods by \( A_1, \ldots, A_n \), where \( A_i \) is the bundle of goods allocated to player \( i \). Spliddit’s algorithm considers three increasingly weaker levels of fairness:

1. **Envy-freeness**: \( V_i(A_i) \geq V_i(A_{i'}) \) for every pair of players \( i, i' \). It is clear that envy-freeness is not always feasible.
2. **Proportionality**: \( V_i(A_i) \geq V_i(G)/n \). In Spliddit, this means each player assigns at least \( 1000/n \) points to her bundle. Again, clearly proportionality may not be feasible.
3. **Maximin share guarantee**: The maximin share (MMS) guarantee of player \( i \) is

\[
\text{MMS}(i) = \max_{X_1, \ldots, X_n} \min_j V_i(X_j),
\]

where the max is taken over partitions of the items into \( n \) subsets \( X_1, \ldots, X_n \). Intuitively, this is the value player \( i \) could guarantee if she divided the items into \( n \) bundles, but then selected a bundle last. An MMS allocation satisfies \( V_i(A_i) \geq \text{MMS}(i) \) for all players \( i \). Although an MMS allocation may not exist [Procaccia and Wang 2014], counterexamples are elaborate and extremely...
unlikely to occur in practice. Moreover, an approximate MMS allocation is guaranteed to exist. Specifically, we can always find an allocation such that \( V_i(A_i) \geq \frac{2}{3} \text{MMS}(i) \) [Procaccia and Wang 2014].

Spliddit’s algorithm works as follows. First, it finds the highest feasible level of fairness. If envy-freeness and proportionality are infeasible, the algorithm computes the maximum \( \alpha > 0 \) such that each player can achieve an \( \alpha \) fraction of her MMS guarantee. Second, the algorithm maximizes social welfare — \( \sum_i V_i(A_i) \) — subject to the fairness constraint found in the first phase. Crucially, while we believe that it will always be possible to find an MMS allocation (with \( \alpha = 1 \)), \( \alpha \geq 2/3 \) is provably feasible even in the worst case [Procaccia and Wang 2014]. This fact enables us to specify an indisputable fairness guarantee, honoring Spliddit’s promise (as made by its tagline) to provide provably fair solutions.

**Sharing credit.** The third and final application divides credit for a joint project between \( n \) players. Possible use cases include determining scientific credit for a paper, dividing a company bonus based on employees’ relative contributions, or sharing credit for a class project. Player \( i \) reports how the rest of the credit should be distributed among the other players. For example, if player 1 thinks that players 2 and 3 contributed equally, and player 4 contributed twice as much, then player 1 would report the contributions 25%, 25%, and 50%, respectively. A solution divides 100% of the full credit for the project among the players.

Work by de Clippel et al. [2008] provides a family of rules for credit division; Spliddit implements one of them.\(^5\) This solution is guaranteed to satisfy a slew of desirable properties. Most importantly, it is impartial: a player’s share of the credit is independent of her report. The solution is also consensual: if there is a division that agrees with all players’ reports, then it is the outcome. While consensuality seems to be quite unrestrictive at first glance, de Clippel et al. [2008] show that if \( n = 3 \), an impartial and consensual rule cannot be exact, that is, it would have to sometimes allocate less than 100% credit overall.\(^6\) Spliddit therefore enforces \( n \geq 4 \).

It is worth noting that one of the possible use cases of the credit division application — sharing scientific credit for a paper — gives rise to interesting questions. Determining the order of authors is a notorious source of acrimony in scientific fields where contribution-based ordering is the norm (that is, almost all scientific fields). Ordering authors by their (fair) share of the credit is an obvious solution. But doing so would not preserve impartiality! Intuitively, while a player cannot change her own share of the credit, she can decrease another player’s share below her own, thereby increasing her position in the author ordering. Very recent work by Berga and Gjorgjiev [2014] establishes impossibility results for impartial ranking under a strong notion of impartiality. Despite this theoretical difficulty, we do believe it is reasonable (albeit not ideal) to use Spliddit’s credit division application for the explicit purpose of ordering authors.

\(^5\)Specifically, the formula in Equation (17) of the paper of de Clippel et al. [2008] is used, with arithmetic means as the aggregators \( \rho \) and \( \tau \).

\(^6\)Note that normalizing the shares would violate impartiality: by changing the sum of shares, a player can change her own normalized share of the credit.
3. IMPLEMENTATION DETAILS

The implementation of Spliddit is centered around a single web application written in Ruby on Rails. All of Spliddit’s nouns, including application instances, players, goods, valuations, and allocations, are modeled using Ruby on Rails Active Record, an object-relational mapping system. On the front end, Spliddit uses the jQuery Javascript library to create interactive user interfaces such as the one used for bidding. Spliddit takes advantage of several services offered by Amazon’s cloud computing platform: Elastic Compute Cloud (for hosting Spliddit’s webservers and running the division algorithms), Elastic Beanstalk (for deploying and automatically scaling Spliddit to handle heavy web traffic), Relational Database Service (for hosting Spliddit’s database), and Simple Email Service (for sending email notifications).

The algorithms described in Section 2 are written in Java for improved performance and ease of implementation. The splitting rent application employs a method of Ünver [2007] for efficiently computing the set of overdemanded rooms based on the Gallai-Edmonds Decomposition Lemma for bipartite graphs. The dividing goods application uses IBM’s CPLEX Optimizer and IBM’s Concert Technology for modeling and solving mixed integer linear programs in Java (mixed integer linear programs are required so that users can mark goods as either divisible or indivisible). Each of the three fairness properties described in Section 2 can be expressed as linear constraints; however, for the weakest level of fairness, the algorithm must first solve mixed integer linear programs to compute each player’s MMS guarantee. The implementation for the sharing credit application is straightforward in comparison, involving the computation of several summations. Since these algorithms can be time consuming when inputs are not very small, Spliddit uses Delayed Job to run the algorithms in background processes, freeing the webservers to quickly respond to incoming HTTP requests.

4. CLOSING REMARKS: EMPIRICAL FAIR DIVISION RESEARCH

As noted above, Spliddit’s two primary goals are providing access to fair division methods, and outreach. However, Spliddit also has the potential to become a revolutionary platform for empirical fair division research. Indeed, as noted by Herreiner and Puppe [2009], fairness properties such as envy-freeness are difficult to study in the lab. A typical experiment in the context of indivisible goods informs each participant of her “value” for each virtual “good”, and pays each participant based on her value for her allocated bundle. In this setting envy-freeness is problematic, because a participant does not truly care about which goods were allocated to another participant (as the goods have no real value) — presumably she mainly cares about how much other participants were paid.

In contrast, Spliddit allows us to partition thousands of users (who report their values for real goods) into multiple groups, and employ a different solution for each group. Happiness surveys will then allow us to gauge the relative importance of various criteria in a way that was previously impossible.

REFERENCES


Display advertisements vary in the extent to which they annoy users. While publishers know the payment they receive to run annoying ads, little is known about the cost such ads incur due to user abandonment. We conducted two experiments to analyze ad features that relate to annoyingness and to put a monetary value on the cost of annoying ads. The first experiment asked users to rate and comment on a large number of ads taken from the Web. This allowed us to establish sets of annoying and innocuous ads for use in the second experiment, in which users were given the opportunity to categorize emails for a per-message wage and quit at any time. Participants were randomly assigned to one of three different pay rates and also randomly assigned to categorize the emails in the presence of no ads, annoying ads, or innocuous ads. Since each email categorization constituted an impression, this design, inspired by Toomim et al. (2011), allowed us to determine how much more one must pay a person to generate the same number of impressions in the presence of annoying ads compared to no ads or innocuous ads.

Categories and Subject Descriptors: J.4 [Social and Behavioral Sciences]: Economics
General Terms: Economics, Experimentation
Additional Key Words and Phrases: display, advertising, quality, compensating differential

1. INTRODUCTION
Display advertising is the prevalent way for publishers to monetize content on the Web. Publishers receive payment from advertisers for placing ads near their content or in their applications. Publishers are typically paid by the number of impressions they can deliver. Thus, they have an incentive to attract and retain users with valuable content, experiences, and applications, and have a disincentive to lose users due to annoyances.

Display ads vary in the extent to which they annoy users. Annoying ads are a source of tension for publishers since they both make money, through payments from advertisers, and cost money, through a decrease in page views due to users abandoning the site. This tension has led to conflict within publishing organizations...
between salespeople, who have an incentive in the form of commission to sell any ads, and management, who are concerned with long-term growth of users and traffic. The continued long-term display of annoying ads may exert negative effects on the publisher, the user, and the advertiser, which we discuss in turn.

First, annoying ads can exert negative effects on publishers. Apart from the user abandonment effects we investigate in this paper, annoying ads might signal that the website, on which the ad is placed, lacks stability (“Why should I trust my email to a site that is so desperate for cash it accepts ads of such poor quality?”), reputability (“Why should I trust the objectivity of a site that is so in the pocket of advertisers it won’t refuse any of them?”), or safety (“Why would I trust this publisher to protect me from phishing attacks, scams, malware, etc. if they are so indiscriminate about who they let advertise?”).

Second, annoying ads can exert a negative impact on users. Ads with excessive animation can get in the way of the user consuming the publisher’s content, undermining the very reason that brought them to the site.

Finally, annoying ads may harm the advertiser that created them. As will be shown, annoying ads are often characterized by exaggerated attempts to capture visual attention such as through fast-moving animation or bizarre imagery. While these manipulations do capture attention, they may also signal that the advertiser is desperate for business or low on resources, undermining the classical signal of quality that advertising is theorized to bring [Riley 2001]. Furthermore, experiments have shown that too much animation can result in lower ad recognition rates compared to ads with moderate or no animation [Yoo and Kim 2005; Burke et al. 2005]. In these ways, annoying ads may actually lower brand reputation and recall, two metrics advertisers typically strive to increase.

If annoying ads exhibit so many negative effects for publishers, users and advertisers, one may wonder why a publisher would run annoying ads at all. The answer may be that it is has been historically difficult to measure the monetary cost of annoying ads. The main contribution of this work is that we measure the compensating wage differential of annoying ads. That is, we measure how much more one must pay a user to do the same amount of work in the presence of annoying ads compared to innocuous ads or no ads. The compensating differential is important to measure because it captures some of the negative effects of advertising, which publishers need to heed as a lower bound when setting the price to run an ad.

Using two experiments, we compute the compensating differential for annoying ads. In the first experiment users rated a set of ads in terms of how annoying they found each ad. In the second experiment, we use those ads identified as more or less annoying, along with the recent methodological innovation of Toomin et al. [2011], to estimate the pay rate increase necessary to generate an equal number of page views in the presence of annoying ads, compared to innocuous ads or no ads. This estimate is the cost of annoying ads in our experiment. We chose categorizing emails as the task to proxy for using a publisher’s site because users either implicitly or explicitly need to categorize their emails as spam or not spam in the presence of ads when using free web-based email services such as Yahoo! Mail, GMail, and Outlook.com.
2. RATING THE QUALITY OF ADS

We now describe the design and results of our first experiment, which served to identify sets of more and less annoying ads (henceforth “bad ads” and “good ads” for brevity) for use in the second experiment.

2.1 Method

The goal of this experiment is to rank a set of actual display ads in terms of annoyingness. After previewing all the ads, users were shown each ad individually, in random order, and asked to rate each ad on a 5-point scale with the following levels: 1) Much less annoying than the average ad in this experiment, 2) A bit less annoying than the average ad in this experiment, 3) Average for this experiment, 4) A bit more annoying than the average ad in this experiment, and 5) Much more annoying than the average ad in this experiment. If we view attributes of an ad as residing in a multidimensional space, the average ratings indicate how users project that multidimensional space onto a one-dimensional annoyingness scale. The 10 most and least annoying ads serve as the sets of “bad” and “good” ads in the next experiment.

3. MEASURING THE COST OF ADS

We use the method of Toomim et al. [2011] along with the sets of “bad” and “good ads” to measure the cost of annoying ads.

3.1 Method

The participants were 1223 Mechanical Turk workers who participated for a base pay of 25 cents and a bonus. Upon accepting the task, participants were randomly assigned to one of nine conditions: three pay conditions and three ad conditions. The pay conditions offered a bonus of one, two, or three cents per five emails classified (i.e., .2, .4, or .6 cents per email), and the ad conditions varied whether “bad ads”, “good ads”, or no ads were displayed in the margin as the task was completed. A chi-squared test found no significant difference in the number of participants beginning work across the nine conditions.

In all conditions, the task consisted of classifying the content of emails as “spam”, “personal”, “work” or “e-commerce” related. Emails were drawn from the public-domain Enron email dataset\(^1\) with one email presented per page, along with accompanying ads, if any. In the “bad ads” condition, two ads randomly drawn from the 10 most annoying ads in our first experiment were displayed in the margins around the email being classified. The “good ads” condition was the same, except the ads were drawn from the 10 least annoying ads. In both conditions, ads were drawn randomly from their respective pools with each page load, and the URLs for the ads were such that ad blocking software would not filter them out. The “no ads” condition simply had whitespace in the margin. The footer included two buttons: one allowing them to submit and rate another email, and a second allowing them to stop categorizing and collect their payment.

\(^1\)http://www.cs.cmu.edu/~enron/ Identifying information such as email addresses, phone numbers and the name “Enron” were removed.
3.2 Results

Let an impression be one participant viewing one email (and its accompanying ads, if any), regardless of whether the participant classifies the email or quits before classifying it. The overall distribution of impressions per person is skewed with a mean of 61, a median of 25 and first and third quartiles of 6 and 57. Being bounded by 1 from below and effectively unbounded from above (only two participants reached the upper limit), these impressions constitute count data. In particular, they are overdispersed count data relative to the Poisson (observed variance / theoretical Poisson data variance is 228.7, $p < .0001$) and thus well suited to a negative binomial generalized linear model (GLM) [Venables and Ripley 2002]. The effects of the conditions on raw impressions are most easily seen in Figure 1, which also makes clear that the difference in impressions between the “good ads” and “no ads” conditions is not significant. Relative to a baseline of “bad ads”, both the “good ads” condition and the no ads condition led to substantially more impressions (19% and 25% more impressions, respectively).

The model expressed in Figure 1 can be used to estimate the compensating differential of annoying ads in this experiment. Since the curves are slightly nonlinear, a range of compensating differentials could be calculated across the pay rate and ad conditions. To get a simple, single approximation we use the middle, “good ads” condition to estimate the effect of pay raises. We take the average of the .2 to .4 and .4 to .6 cent differences, giving an estimated increase of 16.58 impressions resulting from a .2 cent per impression pay raise. When summarizing the effect of ad quality, we use the number of impressions at the .4 cent pay rate. Moving from “bad ads” to no ads, impressions increase by 12.68. The pay raise required to achieve a 12.68 impression increase is .153 cents per impression ($= .2 \times 12.68/16.58$) or $1.53 \text{ CPM}$ (cost per thousand impressions). That is, in this experiment, a participant in the “bad ads” condition would need to be paid an additional $1.53 per thousand
impressions to generate as many impressions as a person in the condition without ads. Similarly, moving from the “bad ads” condition to the “good ads” condition resulted in an additional 9.52 impressions per person. It would require a pay raise of .115 cents per impression (= .2∗9.52/16.58) to generate 9.52 additional impressions, meaning that people in the “bad ads” condition would need to be paid an additional $1.15 CPM to generate as many impressions as in the “good ads” condition.

4. CONCLUSION

The main result of this paper is that annoying ads lead to site abandonment and thus fewer impressions than good ads or no ads. In what might be seen as good news for publishers, good ads and no ads led to roughly equal numbers of impressions.

We calculated the compensating wage differential in our experiment of bad ads to no ads to be $1.53 CPM, and bad ads to good ads to be $1.15. Some care must be taken in interpreting these numbers. While we picked a task—classifying emails—that should be familiar and common for most internet users, this task may not be representative of other internet tasks like reading news stories or searching for products to purchase. Abandonment rates may differ with different tasks and the effects of advertising may vary as well. While virtually every web service features competition, the switching costs vary from very low in consuming news to relatively high in changing email services. Because our users on Mechanical Turk have an outside option of working on an alternative task, we expect our results to be most applicable to situations involving lower switching costs. Nevertheless, we expect that our finding that annoying ads cost the user at least $1 CPM over more pleasant ads will be obtained in some other environments.

For these reasons, we suggest further studies be done on Mechanical Turk, as field experiments, and in laboratories to measure this differential on similar and different tasks. If studies across various domains with a variety of tasks and outside options arrive at similar differentials, more credence can be placed on these numbers. We view this work as a first step in this direction. If future work arrives at similar estimates across a variety of publishers, such estimates could serve as a useful lower bound for what a publisher should charge to run these ads. Moreover, it will be valuable to use the compensating differentials approach to price the various bad aspects of ads, including animation and poor aesthetics.

This work also suggests a variety of policy recommendations. Most directly, the $1 CPM user cost of bad ads has practical consequences for publishers, especially as bad ads often command lower CPMs. It is a reason that publishers should insist on a substantial premium for annoying advertisements. Moreover, a publisher could randomize which users see which ads and track both time spent on the page and the frequency with which users return to the site. This type of experimentation would capture longer term effects of annoying ads than those studied here. Also, publishers could give users an option to close or replace an ad. A replacement event would allow the publisher to infer that a user would prefer a random ad over the ad currently shown. Advertisers with a high closure rate should be charged more since more annoying ads would be closed or replaced faster than less annoying ads. Ad replacement would help the user by removing the annoying ad and the publisher by making it possible to charge for two impressions.
REFERENCES


