Table of Contents

Editor's Introduction

YILING CHEN
Harvard University

I am delighted to introduce Issue 10.3 of SIGecom Exchanges. This issue features one book announcement, seven research letters, a puzzle, and a solution to the *Baffling Raffling* puzzle of Issue 10.1.

The first contribution in this issue is an announcement of the book "Human Computation" by Law and von Ahn. The book is an attempt to better define human computation as a research area, review existing work in this area, draw connections to a wide array of disciplines, and suggest promising research directions.

The seven featured letters cover a variety of topics: mechanism design for information propagation, dynamic auctions, pricing in social networks, security games, incentive-compatible machine learning, prediction markets, and matching theory.

Babaioff, Dobzinski, Oren, and Zohar present mechanisms that incentivize information propagation over a network when nodes receiving the information compete for the same prize. Using Bitcoin, a peer-to-peer electronic currency system, as an example, they characterize reward schemes that incentivize information propagation, are Sybil-proof, and have little payment overhead.

Iyer, Johari, and Sundararajan examine dynamic auctions with learning bidders. They use a mean field equilibrium concept as an approximation method to simplify the analysis, and manage to show the existence of a mean field equilibrium under mild conditions, characterize a simple, optimal strategy for bidders in a mean field equilibrium, and establish a dynamic version of the revenue equivalence theorem.

Candogan, Bimpikis, and Ozdaglar study optimal pricing in social networks for products that exhibit a local network effect. They investigate a monopolist's pricing problem in three different settings — the monopolist can use perfect price discrimination, set a single uniform price, or offer either a full price or a discounted price to each agent.

Vorobeychik investigates security games when security failures are interdependent. He describes both a centralized approach, where an attacker can optimally attack one of the interdependent assets and a centralized defender needs to decide on a security strategy, and a decentralized approach, where defenders make local security decisions while the failures are random.

Meir and Rosenschein report recent advances in strategyproof machine learning, describing the capabilities and limits of strategyproof classification algorithms. In particular, they establish connections between strategyproof classification, facility location, and mechanism design without money.

Ellis and Sami discuss a semester-long experiment on using prediction markets as a classroom learning tool. They report that although no significant improvement in students' enthusiasm or extent of topical reading was found, students reading more broadly at the course start were more likely to trade actively in the prediction market.

Hatfield and Kominers survey recent work in generalized matching theory with a focus on trading networks with transferable utility. They note an interesting parallel between trading networks with discrete contractual opportunities and those with continuously adjustable participation levels. They highlight key conditions — substitutability for the former and concavity for the latter — under which stable outcomes exist, are in the core, and correspond to competitive equilibria.

Finally, there are the puzzles. Our Puzzle Editor, Daniel Reeves, brings the *Contingency Exigency* puzzle. This puzzle asks how to fairly pay an honest worker contingently for a job that has uncertain outcomes. McAfee contributes *Baffling Raffling Debaffled*, a solution to the *Baffling Raffling* puzzle from Issue 10.1. He points out that the puzzle is a special case of a Cournot problem and not only provides a solution but also proves its uniqueness. (No correct solution has been received yet to the *Borrowing in the Limit as Our Nerdiness Goes to Infinity* puzzle from Issue 10.2.)

I would like to thank our Information Director, Felix Fischer, who has once again been very helpful in putting this issue together.

Book Announcement: *Human Computation*

EDITH LAW
Carnegie Mellon University
and
LUIS VON AHN
Carnegie Mellon University

We announce our book, *Human Computation*, published in August 2011 in the Morgan & Claypool Synthesis Series on Artificial Intelligence and Machine Learning. Our announcement is based on the preface of the book.

Categories and Subject Descriptors: I.2.1 [**Artificial Intelligence**]: Applications

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Human Computation, Artificial Intelligence, Human-Computer Interaction, Crowdsourcing

Human computation is a new and evolving research area that centers around harnessing human intelligence to solve computational problems (e.g., image classification, language translation, protein folding) that are beyond the scope of existing Artificial Intelligence (AI) algorithms. With the growth of the Web, human computation systems can now leverage the abilities of an unprecedented number of people to perform complex computation. There are various genres of human computation applications that exist today. Games with a purpose (e.g., the ESP Game) specifically target online gamers who generate useful data (e.g., image tags) while playing an enjoyable game. Crowdsourcing marketplaces (e.g., Amazon Mechanical Turk) are human computation systems that coordinate workers to perform tasks in exchange for monetary rewards. In identity verification tasks, users perform computation in order to gain access to some online content; an example is reCAPTCHA, which leverages millions of users who solve CAPTCHAs every day to correct words in books that optical character recognition (OCR) programs fail to recognize with certainty.

Despite the variety of human computation applications, there exist many common core research issues. How can we design mechanisms for querying human computers such that they are incentivized to generate truthful outputs? What are some techniques for aggregating noisy or complex outputs from multiple human computers in the absence of ground truth? How do we effectively assign tasks to human computers in order to satisfy the objectives of both the system (e.g., quality, budget and time constraints) and the workers (e.g., desire to succeed, to learn, to be entertained)? What classes of computational problems can be efficiently answered using human computation? What are some programming paradigms for designing human computation algorithms? How can human computation systems leverage the joint efforts of both machines and humans?

Authors' addresses: `edith@cmu.edu`, `biglou@gmail.com`

In the past few years, research on human computation has steadily grown, with new works emerging and scattered across conferences and workshops. This book is an attempt to (1) better define human computation as a research area, (2) provide a comprehensive review of existing work, (3) draw connections to a wide variety of disciplines, including AI, Machine Learning, HCI, Mechanism/Market Design and Psychology, and capture their unique perspectives on the core research questions in human computation, and (4) suggest promising research directions in the field.

For many academic and research institutions, this book is free for download at `http://www.morganclaypool.com/toc/aim/1/1`. Based on the materials of this book, we also presented a tutorial at AAAI 2011 entitled "Human Computation: Core Research Questions and State of the Art", the slides for which can be found at `http://humancomputation.com/Tutorial.html`. As this is a rapidly growing field, it is expected that there will be an updated version of the book in the future. We hope that this book will be a useful resource in the years to come for both newcomers and seasoned researchers who are interested in human computation, or more generally, the study of computational systems with humans in the loop.

On Bitcoin and Red Balloons

MOSHE BABAIOFF
Microsoft Research, Silicon Valley
and
SHAHAR DOBZINSKI
Department of Computer Science, Cornell University
and
SIGAL OREN
Department of Computer Science, Cornell University
and
AVIV ZOHAR
Microsoft Research, Silicon Valley

In this letter we present a brief report of our recent research on information distribution mechanisms in networks [Babaioff et al. 2011]. We study scenarios in which all nodes that become aware of the information compete for the same prize, and thus have an incentive *not* to propagate information.

Examples of such scenarios include the 2009 DARPA Network Challenge (finding red balloons), and raffles. We give special attention to one application domain, namely Bitcoin, a decentralized electronic currency system. We propose reward schemes that will remedy an incentives problem in Bitcoin in a Sybil-proof manner, with little payment overhead.

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

General Terms: Algorithms, Economics, Theory

Additional Key Words and Phrases: Bitcoin, Information Propagation, Mechanism Design

1. INTRODUCTION

In 2009, DARPA announced the DARPA Network Challenge, in which participants competed to find ten red weather balloons that were placed at various locations across the United States [DARPA 2009]. Faced with the daunting task of locating balloons spread across a wide geographical area, participating teams attempted to recruit individuals from across the country to help. The winning team from MIT, incentivized balloon hunters by offering them rewards of $2000 per balloon they locate [Pickard et al. 2011]. Recognizing that notifying individuals from all over the US about these rewards is itself a difficult undertaking, the MIT team cleverly offered additional rewards of $1000 to a person who directly recruits a balloon finder, a reward of $500 to his recruiter, and so on. These additional payments created the incentive for participants to spread the word about MIT's offer of rewards and were instrumental in the team's success. In fact, the additional rewards are necessary: each additional balloon hunter competes with the participants in his vicinity, and

Authors' addresses: `moshe@microsoft.com`, `shahar@cs.cornell.edu`, `sigal@cs.cornell.edu`, `avivz@microsoft.com`

reduces their chances of getting the reward for finding a balloon.

MIT's scheme still requires further improvement. As it is, a participant can create a fake identity, invite the fake identity to participate, and use that identity to recruit others. This Sybil attack increases the participant's reward by 50%.[1] Reward schemes should be resistant to such attacks.

A related setting is a raffle, in which people purchase numbered tickets in hopes of winning some luxurious prize. Each ticket has the same probability of winning, and the prize is always allocated. As more tickets are sold, the winning probability of a specific ticket decreases. In this case again, there is a clear tension between the organizer of the raffle, who wants as many people to find out about the raffle, and the participants who have already purchased tickets and want to increase their individual chances of winning. The lesson here is simple, to make raffles more successful participants should be incentivized to spread the word. One example of a raffle already implementing this is Expedia's "FriendTrips" in which the more friends you recruit the bigger your probability of winning.

Our goal is to design reward schemes that incentivize *information propagation* and counter the dis-incentive arising from the competition from other nodes, and are *Sybil proof* while having a *low overhead* (a total reward that is not too high). In particular, we identify the need for such incentives in the Bitcoin protocol, our main example for the rest of this letter. First, we introduce Bitcoin and explain where the incentive problem shows up.

Bitcoin

Bitcoin is a decentralized electronic currency system proposed by Satoshi Nakamoto[2] in 2008 as an alternative to current government-backed currencies [Nakamoto 2008]. Bitcoin has been actively running since 2009, and has been getting a large amount of public attention over the last year. It represents a radical new approach to monetary systems which has appeared in policy discussions and in the popular press. Its cryptographic fundamentals have largely held up even as its usage has become increasingly widespread.

Bitcoin's appeal lies mainly in the ability to quickly transfer money over the internet, and in its relatively low transaction fees.[3] As of November 2011, there are 7.5 million units of currency in circulation (called *Bitcoins*) which are traded at a value of approximately 3 USD per bitcoin.

Bitcoin relies on a peer-to-peer network to verify and authorize all transactions that are performed with the currency. Transactions are cryptographically signed by the owner of the bitcoins that wishes to transfer them, and are sent to nodes in the peer-to-peer network for authorization. Each node in the network is supposed

[1] Indeed, we have no evidence of such attacks in the DARPA challenge. If no such attacks were made, one possible explanation is the short time span of the challenge and its non-commercial, scientific essence. It seems quite plausible that if the challenge is repeated several times such attacks on the MIT reward scheme would become common.

[2] The name Satoshi Nakamoto appears to be an alias. The real identity of Bitcoin's creator remains a mystery.

[3] There are additional properties that some consider as benefits: Bitcoins are not controlled by any government, and its supply will eventually be fixed. Additionally, it offers some degree of anonymity.

to propagate the transaction to its neighbors. Upon receiving a transaction, each node verifies that it is properly signed by the bitcoins' owner, and then tries to "authorize" the transaction by attempting to solve a computationally hard problem (basically inverting a hash function). This authorization process is a key ingredient in maintaining Bitcoin's security (refer to [Nakamoto 2008] for details). Once a node successfully authorizes a transaction, it sends the "proof" (the inverted hash) to all of its neighbors. They in turn, send the "proof" to all of their neighbors and so on. Finally, all nodes in the network "agree" that the transaction has taken place and was authorized.

In compensation for their efforts, nodes are offered a payment in bitcoins for successful authorizations. The system is currently in its initial stages, in which nodes are paid a predetermined amount of bitcoins that are created "out of thin air". This also slowly builds up the bitcoins supply. But Bitcoin's protocol specifies an exponentially decreasing rate of money creation that effectively sets a cap on the total number of bitcoins that will be in circulation. As this payment to nodes is slowly phased out, bitcoin owners that want their transactions approved are supposed to pay fees to the authorizing nodes.

This is where the incentive problem manifests itself. A node in the network has an incentive to keep the knowledge of any transaction that offers a fee for itself, as any other node that becomes aware of the transaction will compete to authorize the transaction first and claim the associated fee. The consequences of such behavior may be devastating: as only a single node in the network works to authorize each transaction, authorization is expected to take a very long time.

We stress that false identities are a prominent concern in Bitcoin. In fact, the Bitcoin protocol is built around the assumption that nodes can create false identities, thus, for a transaction to be approved, nodes that control a majority of the CPU power in the network should accept it, rather than just a majority of the nodes. The latter is vulnerable to Sybil attacks. Therefore any reward scheme for transaction distribution must discourage such attacks.

2. THE MODEL

We present our model for information propagation in Bitcoin's authorization protocol. For a more detailed presentation, refer to [Babaioff et al. 2011].

We assume for simplicity that the network consists of a forest of complete d-ary directed trees, each of them of height H.[4] We model the authorization process of a single transaction in two phases: a distribution phase and a computation phase.

In the beginning of the *distribution phase* the buyer sends the details of the transaction to the t roots of the trees (which we term *seeds*). Each node v that is aware of the transaction can send the information to any of its children. Before sending to any child it can add any number of fake identities. All of v's fake identities are connected to the same set of children. A node can condition its

[4]The intuition for this simplification is that the number of nodes that are aware of the transaction multiplies by some constant for every additional layer that the transaction travels to. A more exact model would be that of a random graph, but this is harder to solve for. In some sense, building the right incentives in the case of trees is harder, as each node monopolizes the flow of information to its descendants.

behavior only on the *length* of the referral chain above it, which can possibly include false identities that were produced by its ancestors.

In the *computation phase* each node that is aware of the transaction tries to authorize it. If there are k such nodes, each of them has the same probability of $\frac{1}{k}$ to authorize it first. We assume that there is a minimal payment for authorization, normalized to 1, which is necessary to motivate the nodes to work on authorizing the transaction.

When a node succeeds in authorizing a transaction we can reward nodes on the chain (starting at some seed) to that node. This chain may contain false identities as well, but cryptographic tools ensure that no node can remove its ancestors from the chain.

3. REWARD SCHEMES

We suggest a rewarding scheme family called the $(\beta, \mathcal{H})$-almost-uniform family. We then combine schemes from this family to create a hybrid scheme that possesses better qualities.

3.1 $(\beta, \mathcal{H})$-Almost-Uniform Schemes

The rewards of schemes in this family are defined as follows: Suppose that a node v has authorized the transaction, and has a chain of l nodes through which it has received the transaction. If $l > \mathcal{H}$ no node is rewarded (so nodes "far" from the seed do not attempt to authorize the transaction). Otherwise, each node in the chain except v gets a reward of β, and v gets a reward of $1 + (\mathcal{H} - l + 1)\beta$.

Given that there are $\Omega(\beta^{-1})$ seeds, the $(\beta, \mathcal{H})$-almost-uniform scheme creates the incentives for each node to propagate information to all its children without duplicating itself. Specifically, we show:

(Informal) Theorem 1. *If there are $\Omega(\beta^{-1})$ seeds, the $(\beta, \mathcal{H})$-almost-uniform scheme guarantees that only strategy profiles that exhibit information propagation and no duplication survive every order of iterated removal of dominated strategies. Furthermore, there exists an order in which no other strategy profiles survive.*

This gives us two interesting schemes, for two different values of β, that offer tradeoffs between the total payment and the number of seeds that need to be initially notified. The first scheme is the $(1, \mathcal{H})$-almost-uniform scheme which requires only a constant number of seeds and its total payment is always $O(\mathcal{H})$. The second scheme is the $(\frac{1}{\mathcal{H}}, \mathcal{H})$-almost-uniform scheme. This scheme works if the number of seeds is $\Omega(\mathcal{H})$. Its total payment is 2.

3.2 The Hybrid Scheme

We combine the $(\frac{1}{\mathcal{H}}, \mathcal{H})$- and $(1, 1 + \log_d \mathcal{H})$-almost-uniform schemes to create a hybrid scheme that requires only a constant number of seeds and pays only a constant amount in expectation. We obtain the following result:

(Informal) Theorem 2. *In the hybrid rewarding scheme, if the number of seeds is at least* 14*, the only strategies that always survive iterated elimination of dominated strategies exhibit information propagation and no duplication. In addition, there exists an elimination order in which the only strategies that survive* *exhibit information propagation and no duplication. Furthermore, the expected sum of payments is at most* 3.

4. DOMINANT STRATEGY MECHANISMS

Iterated removal of dominated strategies is a strong solution concept, but ideally we would like our rewarding scheme to achieve all desired properties in the stronger notion of dominant strategies equilibrium. However, we show that in every dominant strategy scheme either the amount that the scheme must pay in equilibrium is huge, or the number of initial seeds t must be very large.

(Informal) Theorem 3. *Every individually rational reward scheme that propagates information to at least half of the network, and in which no-duplication and information-propagation is a dominant strategy for all nodes, has expected payment of at least* $\frac{1}{10}\left(\frac{2^{H-4}}{t^2}+\frac{1}{t}\cdot\left(\frac{H-3}{t\cdot e}\right)^{H-3}\right)$.

Notice that for the sum of rewards to be constant the number of seeds t has to be a significant part of the network. This implies that dominant strategy schemes are quite impractical.

5. CONCLUSION AND FUTURE RESEARCH

We propose a novel low cost reward scheme that incentivizes information propagation and is Sybil proof. Currently we model the network as a forest of t complete d-ary trees. A challenging open question is to consider the setting where the network is modeled as a random d-regular graph. Other interesting extensions to consider are models that account for the different computation power of nodes, costs of communication, and non-regular graphs (with varying degrees at each node).

REFERENCES

Babaioff, M., Dobzinski, S., Oren, S., and Zohar, A. 2011. On bitcoin and red balloons (full version). Available online: http://research.microsoft.com/apps/pubs/?id=156072q.

DARPA. 2009. The DARPA network challenge. Available online at http://archive.darpa.mil/networkchallenge/.

Nakamoto, S. 2008. Bitcoin: A peer-to-peer electronic cash system. Available online at http://bitcoin.org/bitcoin.pdf.

Pickard, G., Pan, W., Rahwan, I., Cebrián, M., Crane, R., and Madan, A. 2011. Time critical social mobilization. *Science 334*, 509–512.

Mean Field Equilibria of Dynamic Auctions with Learning

KRISHNAMURTHY IYER and RAMESH JOHARI
Stanford University
and
MUKUND SUNDARARAJAN
Google Inc.

We study learning in a dynamic setting where identical copies of a good are sold over time through a sequence of second price auctions. Each agent in the market has an *unknown* independent private valuation which determines the distribution of the reward she obtains from the good; for example, in sponsored search settings, advertisers may initially be unsure of the value of a click. Though the induced dynamic game is complex, we simplify analysis of the market using an approximation methodology known as *mean field equilibrium* (MFE). The methodology assumes that agents optimize only with respect to long run average estimates of the distribution of other players' bids. We show a remarkable fact: in a mean field equilibrium, the agent has an optimal strategy where she bids truthfully according to a *conjoint valuation*. The conjoint valuation is the sum of her current expected valuation, together with an overbid amount that is exactly the expected marginal benefit to one additional observation about her true private valuation. Under mild conditions on the model, we show that an MFE exists, and that it is a good approximation to a *rational* agent's behavior as the number of agents increases. We conclude by discussing the implications of the auction format and design on the auctioneer's revenue. In particular, we establish a dynamic version of the revenue equivalence theorem, and discuss optimal selection of reserve prices in dynamic auctions.

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: —*Economics*

General Terms: Economics, Theory

Additional Key Words and Phrases: Dynamic auctions, Learning, Mean field equilibrium

1. INTRODUCTION

Auctions are observed as a market mechanism in a wide range of economic transactions: sponsored search markets run by Google and Yahoo!, online marketplaces such as eBay and Amazon, crowdsourcing, procurement auctions for public service contracts, licensing auctions (e.g., for mining or oil tracts), etc. Nearly all of these examples are characterized by two important features. On one hand, the auction format is typically relatively straightforward to describe, consisting of repetitions of a simple one-shot auction format. On the other hand, despite the simplicity of the mechanism itself, such markets can give rise to complex dynamic incentives for bidders. As a result, many basic questions become quite challenging: determining optimal bidding strategies for bidders; characterizing dynamic equilibrium behavior among the bidders; and determining optimal choices of market parameters for the auctioneer, such as auction format and reserve prices.

Authors' addresses: `kriyer@stanford.edu`, `ramesh.johari@stanford.edu`, `mukunds@google.com`

As a concrete example, consider online sponsored search auctions [Edelman et al. 2007; Varian 2007]. These repeated auctions operate on a per keyword basis; in a typical scenario, advertisers bidding on a particular keyword estimate their underlying valuation based on the conversion rate from an ad click to revenue (e.g., from sales). As the advertisers win more ad placement through the auction, they *learn* this conversion rate, which informs their bidding decision in the auction [Ghose and Yang 2009; Rey and Kannan 2010; Sculley et al. 2009]. In these settings, the bidders face a trade-off between *exploration*, where they bid higher to obtain more information about their value, and *exploitation*, where they bid optimally given their current information.

The exploration-exploitation trade-off has significant ramifications for market operation and design. *First*, it complicates the design of optimal bidding strategies. In sponsored search markets, for example, identifying optimal learning strategies would lead to better design of bidding agents that incorporate advertisers' uncertainty about their conversion rate. *Second*, as the bidders are playing a complex dynamic game, it can be intractable to characterize equilibrium behavior among many interacting bidders, and in particular to determine how bidders' uncertainty affects the distribution of bids seen over time. *Third*, as a consequence, we lose the ability to guide market operation and design. For instance, auctioneers usually set reserve price in such markets to increase their revenue. As we later demonstrate, setting a reserve without incorporating the learning among the bidders may cause unwarranted restriction of allocation and ultimately yield *lower* revenue.

In our paper [Iyer et al. 2011], we study an abstract dynamic auction model that consists of repetitions of a simple one-shot mechanism. We primarily study a setting where identical copies of a good are sold through a sequence of second price auctions over time; as an example, a copy of the good may denote a click on an advertiser's ad in sponsored search settings. (We also analyze repetitions of other standard one-shot auction formats.) Each agent in the market has an independent private valuation that determines the distribution of the reward she obtains from the good; the private valuation may denote an advertiser's conversion rate in sponsored search. Although agents are initially unaware of their own private valuation, every time an agent wins an auction and obtains a copy of the good, her realized reward from the good incrementally informs her about her valuation. The strategic interactions among the agents along with their beliefs about their valuation influence their bids in the auction. Thus, we naturally obtain a dynamic game among the agents in our model.

The standard game-theoretic tool used to analyze such dynamic games is the equilibrium concept known as perfect Bayesian equilibrium (PBE). However, there are two central problems with this approach. First is that such equilibria are *intractable*: the state space complexity is enormous (since bidders must maintain beliefs over all that is unknown to them), and grows exponentially with the number of bidders and with time. Second, and partly in consequence, is that such equilibria are *implausible*: in equilibrium, a bidder's optimal bidding strategy is intricately predicated on what she believes *other* bidders' strategies are, and a PBE requires each bidder to accurately forecast and estimate *exactly* how her competitors will respond to any bid she makes today.

The complexity of PBE motivates us to consider an approximation methodology that we refer to as *mean field equilibrium* (MFE). (See [Adlakha et al. 2010; Huang et al. 2007; Jovanovic and Rosenthal 1988; Lasry and Lions 2007; Tembine et al. 2009].) MFE is inspired by a *large market* approximation: with a large number of bidders, tracking and forecasting the exact behavior of individual bidders is impractical and implausible. In an MFE, individuals take a simpler view of the world. They postulate that fluctuations in the empirical distribution of other agents' bids have "averaged out", and thus optimize holding the bid distribution of other agents fixed. MFE requires a consistency check: the postulated bid distribution must arise from the optimal strategies agents compute. The benefit of analyzing a large market using MFE is that for the agents to optimize their behavior, it is sufficient for them to just maintain beliefs about their *own* private valuation. This reduces the dimension of the system state that each agent needs to track, simplifying the analysis tremendously.

Furthermore, we believe that MFE corresponds more closely to an equilibrium concept that might be applicable in practice, particularly in settings with a large number of bidders. For example, in sponsored search auction markets, bidders generally do not have access to complete information about the bid history for auctions they participated in. Rather, bidders are usually provided with various tools by the auctioneer to aid in strategizing how to bid; e.g., Google provides the advertisers with a bid simulator that simulates how often an ad would get displayed and clicked upon on making a particular bid [Friedman 2009]. The bid simulator bases its predictions on aggregated historical data, that gives a "bid landscape" of competitors' bids on the same category or keyword (i.e., the distribution of bids). Bid landscapes inherently assume stationarity in the market, at least for a limited time horizon of interest; thus bidders are reacting to average information about their competitors. It is reasonable to expect that for many bidders, therefore, their own decision of how to bid will *not* explicitly forecast opponents' reactions, and instead will assume that these reactions have averaged out in any forecasting about future auction outcomes. This type of example illustrates how the rationality assumptions in MFE might naturally arise in practice.

Our main contributions address the challenges raised above.

(1) *Characterizing optimal strategies for bidders: Conjoint valuations.* We show that in the large market model, the optimal strategy of an agent takes a remarkably simple form: given her current belief about her valuation, the optimal strategy is to bid according to a *conjoint valuation.* The conjoint valuation is the sum of her current expected valuation, together with an overbid. This overbid denotes an agent's value for learning about her true private valuation, and we show that it is exactly the expected marginal benefit to one additional observation about her valuation. Thus the conjoint valuation presents a structurally simple and *plausible* strategy that captures how an agent in the large market balances the trade-off between exploration and exploitation.

(2) *Consistency and validity of the mean field model: Existence of MFE and an approximation theorem.* We show that the mean field model is consistent by proving the existence of an MFE. This involves showing that the stationary distribution of a market where each agent follows the mean field strategy turns

out to be the market distribution that each agent had assumed to solve their decision problem. We extend this result under mild conditions to a dynamic auction setting consisting of repetitions of a fixed *standard auction*; for example, standard auctions include second price, first price, and all-pay auctions. Thus, we obtain, in fairly general settings, the existence of informationally simple equilibria in mean field models, where agents make bids taking into consideration only their own belief about their valuation and the bid distribution in the market. This result provides evidence of the *tractability* of the mean field model.

We next tackle the issue of whether an MFE, which rests on a large market assumption, accurately captures a rational agent's behavior in a finite market. We prove that indeed an MFE is asymptotically a good approximation to agent behavior in a finite market. Formally, we show that if in a finite market, every agent except one follows the MFE strategy, then the remaining agent's loss on playing the MFE strategy converges to zero as the number of agents in the market increases. This result justifies formally the use of an MFE to analyze agent behavior in a finite market as the number of agents increases. We emphasize, however, that MFE may be a useful approximation even when the number of agents is not large, simply because it more accurately captures the information available to bidders when they optimize (e.g., in sponsored search auctions, bidders are responding to bid landscape information).

(3) *Market design: Auction format and reserve prices.* Finally, to illustrate the power of our approach, we leverage the analytical and computational simplicity of MFE to address "second best" market design: how should an auctioneer choose the auction format to maximize revenue, with the constraint of relatively "simple" repeated auction mechanisms?

In static settings, the revenue equivalence theorem states that an auctioneer's expected revenue in any standard auction remains the same. In a dynamic setting, the main difference is that now changing the auction format not only affects an agent's payment in each auction, but also affects her incentive to learn more about her valuation. Nevertheless, we prove a *dynamic revenue equivalence* theorem that extends the static version to dynamic settings. We show this by relating, under some conditions, an MFE of a market with repetitions of a given standard auction format to an MFE of a market with repeated second price auction. This result shows that changing the one-shot auction format will not increase the seller's expected revenue.

We then consider the possibility of increasing revenue by choosing a reserve price. In static auctions, setting a reserve has the effect of extracting greater revenues from high valuation bidders, at the expense of shutting out bidders with lower valuations [Myerson 1981]. In dynamic auctions with learning, however, a reserve has an added effect: it reduces bidders' incentives to learn their valuation. Ignoring this added effect while setting the reserve may cause the auctioneer to incur a high penalty. We develop benchmarks to evaluate this penalty, by comparing the MFE where an auctioneer anticipates bidders' learning behavior, against one where the auctioneer is oblivious to bidders' learning. The computational tractability of MFE allows us to evaluate these benchmarks:

we numerically observe that depending on the uncertainty bidders hold about their valuations, the incremental benefit to setting a reserve for an auctioneer could be as high as 15-30% more than the incremental benefit if the learning is ignored.

REFERENCES

Adlakha, S., Johari, R., and Weintraub, G. Y. 2010. Equilibria of dynamic games with many players: Existence, approximation, and market structure. *arXiv abs/1011.5537*.

Edelman, B., Ostrovsky, M., and Schwarz, M. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *The American Economic Review 97*, 242–259(18).

Friedman, D. 2009. Bid like a pro with the bid simulator. `http://adwords.blogspot.com/2009/08/bid-like-pro-with-bid-simulator.html`.

Ghose, A. and Yang, S. 2009. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science 55,* 10, 1605–1622.

Huang, M., Caines, P. E., and Malhamé, R. P. 2007. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Transactions on Automatic Control 52,* 9, 1560–1571.

Iyer, K., Johari, R., and Sundararajan, M. 2011. Mean field equilibria of dynamic auctions with learning. *SSRN eLibrary* 1799085.

Jovanovic, B. and Rosenthal, R. W. 1988. Anonymous sequential games. *Journal of Mathematical Economics 17*, 77–87.

Lasry, J. M. and Lions, P. L. 2007. Mean field games. *Japanese Journal of Mathematics 2*, 229–260.

Myerson, R. B. 1981. Optimal auction design. *Mathematics of Operations Research 6,* 1, 58–73.

Rey, B. and Kannan, A. 2010. Conversion rate based bid adjustment for sponsored search. In *Proceedings of WWW*. 1173–1174.

Sculley, D., Malkin, R. G., Basu, S., and Bayardo, R. J. 2009. Predicting bounce rates in sponsored search advertisements. In *Proceedings of ACM SIGKDD*. 1325–1334.

Tembine, H., Boudec, J.-Y. L., El-Azouzi, R., and Altman, E. 2009. Mean field asymptotics of Markov decision evolutionary games and teams. In *Proceedings of GameNets '09*. 140–150.

Varian, H. R. 2007. Position auctions. *International Journal of Industrial Organization 25,* 6, 1163–1178.

Optimal Pricing in Social Networks (Extended Abstract)

OZAN CANDOGAN
Laboratory for Information and Decision Systems, MIT
and
KOSTAS BIMPIKIS
Graduate School of Business, Stanford University
and
ASUMAN OZDAGLAR
Laboratory for Information and Decision Systems, MIT

We consider the pricing strategies of a monopolist selling a divisible good (service) to consumers who are embedded in a social network. We assume that each consumer's usage level depends directly on the usage of her neighbors in the social network, and investigate the optimal pricing policies of the monopolist. We show that if the monopolist can perfectly price discriminate the agents, then the price offered to each agent has three components: a nominal price, a discount term due to the agent's influence on her neighbors, and a markup term due to the influence of her neighbors on the agent. We also characterize the optimal pricing strategies in settings where the monopolist is constrained to offering a single price, and where she can choose two distinct prices (a discounted and a full price). For the former setting we provide a polynomial time algorithm for the solution of the pricing problem. On the other hand, we show that in the latter setting the optimal pricing problem is NP-hard, and we provide an approximation algorithm, which, under some conditions, achieves at least 88% of the maximum profit.

Categories and Subject Descriptors: K.4.4 [**Computers and Society**]: Electronic Commerce

General Terms: Economics, Algorithms, Theory.

Additional Key Words and Phrases: Optimal pricing, social networks, externalities.

1. INTRODUCTION

Social networks underpin most human interactions and have traditionally been the main medium through which individuals obtain their information and form their opinions. The torrent of new communication technologies have transformed their structure and increased our reliance on them for social interactions and access to global information. Social, business and political decisions are now, to an unparalleled extent, shaped by information provided by the networks in which they are situated. This area presents a unique opportunity for academic research mainly for two reasons. First, social networks play a central role in a variety of real world environments. For example, it is well documented that they are an integral component of the labor market (see [Montgomery 1991], [Granovetter 1974]) and they are crucial for the successful marketing of consumer products and the adoption of new technologies (see [Ellison and Fudenberg 1993]).

Authors' addresses: candogan@mit.edu, kostasb@stanford.edu, asuman@mit.edu.

Second, due to the overwhelming success of online social networking communities, such as Facebook and Twitter, and the ubiquity of the Internet based services that are built around social networks, there is a wealth of data available on the interaction between individuals and its influence on their decision making. So far, this data has been used mostly as part of research efforts descriptive in nature. Specifically, the focus has been on understanding the data and introducing models on how individuals interact with each other. However, we feel that there is a need for prescriptive research, i.e., research that mainly focuses on the following broad question: how can we use our understanding of the structure and function of social networks to improve outcomes? In our work [Candogan et al. 2011], we take a prescriptive point of view, and investigate whether and how a monopolist that possesses knowledge regarding the social network of her consumers, can derive a pricing policy that uses this information and maximize profits.

More concretely, we focus on products that exhibit a local (positive) network effect: increasing the usage level of a consumer has a positive impact on the usage levels of her peers. Concrete examples of such goods include online games (e.g., World of Warcraft, Second Life) and social networking tools and communities (e.g., online dating services, employment websites etc.). More generally, the local network effect can capture the word of mouth communication among agents. In this setting, we consider a monopolist, who has access to data on network interactions of individuals, and investigate how she could improve her pricing strategies using the relevant data.

2. MAIN RESULTS

We investigate the pricing problem of the monopolist in three different settings. In the first setting, we assume that the monopolist can use perfect price discrimination, i.e., it can offer a different unit price to each of the individuals in the network. In this setting, we show that the optimal prices admit an interesting decomposition to three components: (i) a nominal term which is independent of the network structure, (ii) a discount term proportional to the influence that this agent exerts over the rest of the social network, (iii) and a markup term proportional to the influence that the network exerts on the agent. Both the markup and the discount are proportional to the Bonacich centrality of the agent's neighbors in the social network structure, which is a measure of network influence introduced and used widely in the sociology literature. Thus, our work provides a micro-founded model that motivates use of Bonacich centrality as a measure of influence. Informally, our result suggests that agents get a discount proportional to the amount they influence their peers to purchase the product, and they receive a markup if they are strongly influenced by other agents in the network.

Perfect price differentiation is typically hard to implement. Thus, we also discuss the optimal strategy of a monopolist who can offer a single uniform price for the good. Since the monopolist can use a single price, at the optimal price some of the agents may prefer not to purchase the product. In this setting, the monopolist should choose the price such that price is low enough to induce high total consumption, and large enough to result in large revenues. We develop an algorithm that finds the optimal single price in time polynomial in the number of agents. The algorithm relies on a ranking of agents according to a novel measure of centrality in the underlying network. Using this ranking, the algorithm considers different subsets of the consumers and finds the optimal price provided that only the consumers in the given subset purchase a positive amount of the good. We show that it is sufficient to consider N such subsets (for a network of N agents), and the algorithm computes the optimal price in time polynomial in N.

Third, we consider an intermediate setting, where the monopolist can offer a full price or a discounted price to each agent. In this setting, the decision problem of the monopolist involves deciding which subset of agents should receive the discounted prices. This problem is related to the well-known MAX-CUT problem, which involves finding a graph cut that maximizes the total weight of the edges between the two partitions, obtained as a result of the cut. Exploiting this relation, we establish that the optimal pricing problem is NP-hard. In this setting, we provide an approximation algorithm that recovers (in polynomial time) at least 88 % of the optimal revenue for the case of two prices.

Finally, we study the impact of the availability of network information on monopolist's profits. In particular, we compare profits of a monopolist that does not have access to this information to those of a monopolist that uses this information optimally. We show that when the influence structure is asymmetric (when agent i influences agent j, significantly more than j influences i), the monopolist can significantly improve her profits by using pricing rules which exploit the underlying network structure.

3. CONCLUSIONS

Our recent work [Candogan et al. 2011] shows that firms can significantly improve their pricing decisions by using the available social network data. In particular, we characterize the optimal pricing strategies of a monopolist who has access to the social network information of her consumers. This work leads to a number of interesting future questions. One interesting direction is to see how the optimal pricing strategies change in presence of competition between firms. A second interesting direction is characterizing the pricing strategies in dynamic settings, where the players learn their preferences or the underlying quality of the good. More generally, we feel that the field is in a great need for more prescriptive research that mostly focuses on how to use our, mature by now, understanding of the structure of social networks to improve managerial decision making and policy recommendations. We believe that this is both an extremely challenging and rewarding area for future research.

REFERENCES

Candogan, O., Bimpikis, K., and Ozdaglar, A. 2011. Optimal pricing in networks with externalities. *Submitted*.

Ellison, G. and Fudenberg, D. 1993. Rules of thumb for social learning. *Journal of Political Economy*.

Granovetter, M. 1974. Getting a job: A study of contacts and careers. *Harvard University Press*.

Montgomery, J. 1991. Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review*.

Security and Network Effects: Centralized and Decentralized Perspectives

YEVGENIY VOROBEYCHIK
Sandia National Laboratories, Livermore, CA[1]

Security, like many other complex decisions, is generally approached with a divide-and-conquer mindset. Consequences of security failures, however, can rarely be completely localized: an explosion or a fire at one building can affect neighboring structures, a debt crisis in Greece resonates throughout the tightly connected European and US financial markets, and a breach of security at one computer can facilitate access to others on the same network. It is thus crucial to view security holistically, and devise security strategies that explicitly account for interdependencies between valuable assets. Here we provide an overview of two recent approaches to security with network effects. The first approach takes a centralized perspective, attempting to compute an optimal security configuration for all interdependent assets. This approach explicitly accounts for an intelligent adversary optimally attacking one of the assets. The second approach studies the impact of decentralized decision making when local failures can propagate in complex ways through the entire system, but assumes that initial failures are random.

Categories and Subject Descriptors: I.2.11 [**Artificial Intelligence**]: Distributed artificial intelligence—*Intelligent agents*

General Terms: Algorithms, Performance, Economics, Security

Additional Key Words and Phrases: Game theory, Security, Stackelberg Games, Networks

1. MOTIVATION

Security-related decisions usually take place in a complex, dynamic, highly interdependent environment. Typical mathematical modeling of security naturally abstracts away most of the complexity. In particular, a common simplification is to ignore interdependencies between assets to secure, or parties that make security decisions. Recently, Kunreuther and Heal proposed a rather influential model that captures, in a highly stylized way, interdependencies between security decisions of different parties [Kunreuther and Heal 2003]. While Kunreuther and Heal (and follow-up work) explicitly account for interdependencies between security decisions of different players, they only consider a binary action space for each player: to secure, or not. Thus, they abstract away an important aspect of security: each decision maker is often responsible for securing a collection of interdependent assets. Each individual decision, in isolation of game theoretic interactions, is one rife with complexities, while the strategic aspect adds yet another dimension.

Below I describe two recent approaches that explicitly capture the complexity of interdependent security decisions, both at the individual (centralized) level, and accounting for strategic behavior of multiple self-interested parties.

[1]Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

2. A CENTRALIZED APPROACH TO INTERDEPENDENT SECURITY

In this section we describe a centralized approach to interdependent security by Letchford and Vorobeychik [2011]. Consider a collection of valuable assets (targets), T, which must be defended from a rational attacker. The defender's options involve security configurations for each target (e.g., firewall settings), and a tradeoff between highly effective, but costly, security options (such as a highly restrictive firewall setting that prevents valuable operations from being performed), and cheaper security options which are more fragile. An attacker is endowed with a power to observe the defender's decision, and then execute a single attack against an asset which yields the greatest gain. The defender, however, has an important ally: randomization. Specifically, the defender can "commit" to a randomized strategy by which security configurations are chosen, and the attacker can only observe the stochastic security configuration, but not its ultimate realizations.

While the model described so far is similar to a typical Stackelberg security game setting [Conitzer and Sandholm 2006; Paruchuri et al. 2008; Kiekintveld et al. 2009], Letchford and Vorobeychik [2011] extend this line of work by explicitly representing interdependencies between assets as a graph, $G = (T, E)$, where assets are identified with nodes in the graph, and the set of edges E represents interdependencies between them. Moreover, each asset t has an *intrinsic* value w_t, which is lost to the defender if this asset, and only this asset, is compromised by the attacker (symmetrically, the attacker would gain an intrinsic value v_t from compromising or destroying t). Interdependencies are modeled as independent failure cascades: if an asset t fails (is successfully attacked), its network neighbors t' also fail, independently, each with probability $p_{t,t'}$.

An important assumption in past work on Stackelberg security games has been that security decisions are independent across assets. This assumption is clearly violated in the most general incarnation of the model just described. However, under the assumption that the cascade probabilities do not depend on security configurations, we can attain *effective independence* by simulating failure cascades initiated at each asset in T and quantifying the resulting expected utilities to the defender and attacker. Stated more generally, the requisite assumption is that security decisions targeting external threats are independent of security decisions targeting insider threats (e.g., security threats originating from other computers on a local network), a state of affairs that seems ubiquitous in network security settings. Significantly, we can subsequently use a linear programming approach to compute optimal randomized security policies that accounts for cascading failures.

Letchford and Vorobeychik perform a computational analysis of several interdependent security settings. They observe that total defense expenditures exhibit a single peak as a function of cost in graphs with a relatively homogeneous degree distribution, but two peaks in scale-free graphs. They additionally demonstrate that resilience properties of graphs have significantly different characteristics from those observed when defense decisions are not taken into account.

3. ROBUSTNESS, FRAGILITY, AND DECENTRALIZATION

The approach described in the previous section assumes that the defense decision is entirely centralized. A complementary approach in Vorobeychik *et al.* [2011]

presents a model of security decisions that accounts for interactions between security decisions of multiple players, although this alternative model presumes that the failures happen according to a fixed distribution, rather than as targeted attacks. The Vorobeychik *et al.* model also aims to gain fundamental insights about security in complex systems, rather than to provide a framework that can capture realistic interdependent security settings. As such, the complex interdependencies between security decisions arise based on a very abstract and simplified model of forest fires. The forest fire model starts with a square ($N \times N$) grid. Each cell in this grid is encoded by a binary value, where 1 indicates a presence of a tree in that cell. The grid is partitioned among a set of players, each deciding which cells, among those he owns, are to contain a tree; for every tree a player plants he pays a fixed cost c. The catch is that after a joint decision is made by all players, a lightning can strike any cell in the grid according to a predefined spacial distribution, burning down the entire connected component to which this cell belongs in the process. The goal of each player is to maximize the total *yield*, or expected number of trees he plants that survive the lightning, less total planting costs.

The complexity of this game theoretic forest fire model precludes mathematical analysis of an arbitrary instantiation. However, it is not difficult to obtain insight in the boundary cases where there is either a single player controlling the entire grid, or where each player controls a single cell. In the former case, clearly, socially optimal solution exactly matches the "equilibrium" configuration. In the latter, Vorobeychik *et al.* show that equilibrium solutions can be arbitrarily poor. Simulation-based game theoretic analysis reveals an interesting pattern when the number of players is between these extremes: there is a level of moderate decentralization where equilibrium configurations are, on average, close to socially optimal, and additionally exhibit high resilience to changes in the lightning distribution. In contrast, highly centralized solutions are extremely fragile to environment changes. The high-level reason why decentralization yields greater resilience is that individually, players' decisions must account both for the negative events that impact them directly, as well as the spread of fire due to the selfish choices of their neighbors. Thus, players build in extra robustness into their configurations that is absent in a highly centralized decision.

REFERENCES

Conitzer, V. and Sandholm, T. 2006. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM conference on Electronic commerce*. EC '06. ACM, New York, NY, USA, 82–90.

Kiekintveld, C., Jain, M., Tsai, J., Pita, J., Ordóñez, F., and Tambe, M. 2009. Computing optimal randomized resource allocations for massive security games. In *In AAMAS-09*.

Kunreuther, H. and Heal, G. 2003. Interdependent security. *Journal of Risk and Uncertainty 26,* 2-3, 231–249.

Letchford, J. and Vorobeychik, Y. 2011. Computing randomized security strategies in networked domains. In *AARM Workshop*.

Paruchuri, P., Pearce, J. P., Marecki, J., Tambe, M., Ordonez, F., and Kraus, S. 2008. Playing games with security: An efficient exact algorithm for Bayesian Stackelberg games. In *Seventh International Conference on Autonomous Agents and Multiagent Systems*. 895–902.

Vorobeychik, Y., Mayo, J. R., Armstrong, R. C., and Ruthruf, J. R. 2011. Noncooperatively optimized tolerance: Decentralized strategic optimization in complex systems. *Physical Review Letters 107,* 10, 108702.

Strategyproof Classification

RESHEF MEIR and JEFFREY S. ROSENSCHEIN
The Hebrew University of Jerusalem

Experts reporting the labels used by a learning algorithm cannot always be assumed to be truthful. We describe recent advances in the design and analysis of strategyproof mechanisms for binary classification, and their relation to other mechanism design problems.

Categories and Subject Descriptors: I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent Systems*

General Terms: Algorithms, Theory, Economics

Additional Key Words and Phrases: Mechanism design, Classification, Game theory

1. INTRODUCTION

The field of mechanism design deals with problems that involve multiple parties, or agents, with potentially conflicting interests. The goal is typically to design interaction rules such that rational behavior of the agents will lead to an outcome that is "good" according to a certain criterion. Such a criterion may be the welfare of the agents themselves, or some other achievement about which the designer cares.

Most machine learning problems do not fall into that category. Prior research has traditionally addressed many issues related to the quality of learning (such as noise, biased sampling, partial information, and even multiple experts), but the issue of incentives has received much less attention.

However, when multiple experts are involved, game-theoretic considerations become increasingly important, especially when the agents (i.e., experts) have a direct interest in the outcome of the learning algorithm. More specifically, agents may *lie* so as to bias the outcome closer to their own opinion.

In a SIGecom Exchanges letter a few years ago, Ariel Procaccia [2008] reviewed several *strategyproof learning mechanisms*—that is, learning mechanisms[1] in which agents have no incentive to lie. Unfortunately such mechanisms cannot guarantee an optimal result (in terms of the minimal total error), and thus we evaluate them according to their worst-case approximation ratio, when compared to the optimal outcome.

Procaccia presented truthful approximation mechanisms in two highly important supervised learning domains, namely *regression* and *binary classification*. His letter also called for a synthesis of mechanism design and machine learning, and predicted that such a joint approach will benefit both communities.

In the three years that have passed since the aforementioned letter, both fields have advanced significantly. Furthermore, it turns out that strategyproof learning

[1]We use the term *mechanism* as a higher abstraction level than that of an algorithm; a mechanism focuses on information passed, and incentives, rather than on implementation details.

Authors' addresses: `reshef24@cs.huji.ac.il`, `jeff@cs.huji.ac.il`

does not have to be treated as a standalone mechanism design problem, but that it is deeply related to other kinds of problems as well. In this paper, we describe recent advances in strategyproof classification, and explain some of its unexpected connections to the problems of *facility location* and *judgment aggregation*.

2. STRATEGYPROOF CLASSIFICATION MECHANISMS

We begin with some formal definitions. A *classifier* or *concept* c is a function from some input space $\mathcal{X}$ to *labels* $\{-,+\}$. A *concept class* $\mathcal{C}$ is a set of concepts. Each agent $i \in I$ controls a set of data points X_i, where $Y_i : X_i \to \{+,-\}$ reflects the true label of each data point (known only to agent i). Let $S_i = \{\langle x, Y_i(x)\rangle \ : \ x \in X_i\}$ be the partial *dataset* of agent i, and let $S = \langle S_1, \dots, S_n\rangle$ denote the complete *dataset*.

In a classification problem, we are given a dataset S and a concept class C, and need to return some $c \in C$ which best classifies the data. To evaluate a classifier, we simply count the number of errors (the 0–1 loss). That is, $\mathrm{R}_i(c, S) = \sum_{x \in X_i} [\![c(x) \neq Y_i(x)]\!]$ (where $[\![A]\!] = 1$ iff A is true and 0 otherwise). The *global risk* is similarly defined as $\mathrm{R}_I(c, S) = \sum_{i\in I} w_i \mathrm{R}_i(c, S)$, where w_i is the weight of agent i.

A *classification mechanism* $\mathbf{M}$ is a function (deterministic or randomized) mapping each dataset S to a classifier $c \in \mathcal{C}$; we do not allow a mechanism to make payments. We say that a mechanism is *strategyproof* (SP) if no agent can gain by lying. Formally, if for every S, i and S'_i (where the labels in S_i, S'_i may differ) it holds that $\mathrm{R}_i(\mathbf{M}(S), S) \leq \mathrm{R}_i(\mathbf{M}(S'_i, S_{-i}), S)$.

The classifier with the lowest global risk is called the *empirical risk minimizer* (ERM), and is denoted by $c^*(S)$. The optimal risk is denoted by $\mathbf{opt}(S) = \mathrm{R}_I(c^*(S), S)$. Finally, $\mathbf{M}$ is an α*-approximation* mechanism if for every dataset S, $\mathrm{R}_I(\mathbf{M}(S), S) \leq \alpha \cdot \mathbf{opt}(S)$.

The goal of the strategyproof classification agenda is the design and analysis of SP mechanisms with good (i.e., low) approximation ratios.

The setting reported by Procaccia [2008] (originally published in [Meir et al. 2008]) was a very simple one, and assumed that there are only two possible classifiers, i.e., that $|\mathcal{C}| = 2$. Under this extreme limitation, the authors provided a deterministic 3-approximation SP mechanism, and showed that no better mechanisms exist. Allowing randomization can improve the approximation to 2, which is again a tight bound.

When considering the most general classification setting, no deterministic SP mechanism can guarantee any reasonable outcome [Meir et al. 2010] in terms of approximation.[2] This result holds for widely used concept classes like linear classifiers and boolean conjunctions. This strong negative result means that some restriction is necessary to obtain good mechanisms.

2.1 Shared inputs

A natural restriction is to assume that all agents are labeling the same set of data points X, i.e., that all X_i are equal. This is the case, for example, in online surveys, where everyone is answering the same set of questions. Quite interestingly, this simple restriction makes the problem much easier to handle. In fact, selecting

[2] A similar negative result was provided for randomized mechanisms, but it requires additional technical assumptions.

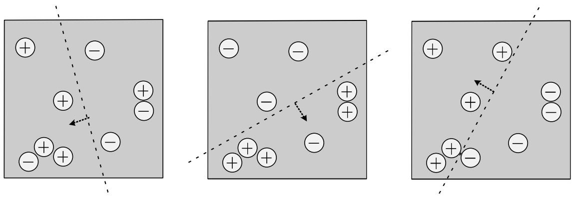

Fig. 1. An instance of a dataset with shared inputs. Here, $\mathcal{X} = \mathbb{R}^2$, $\mathcal{C}$ is the class of linear separators over $\mathbb{R}^2$, and $n = 3$. The data points X of all three agents are identical, but the labels are different. The best classifier from $\mathcal{C}$ with respect to each S_i is also shown (the arrow marks the positive halfspace). Only the rightmost dataset is realizable.

an agent at random and using him as a dictator guarantees a $3 - \frac{2}{n}$ approximation ratio (and is clearly SP) [Meir et al. 2009]. When agents are weighted, the same approximation ratio can be achieved with a more subtle randomization [Meir et al. 2011]. The latter paper also proves that no SP mechanism can do better.

2.2 Realizable datasets

A particular case of interest is when the dataset of an agent can be classified perfectly, i.e., there is some $c_i \in \mathcal{C}$ s.t. $c_i(x) = Y_i(x)$ for all $x \in X$. If this is the case for every agent, we say that the data is *individually realizable* (IR); see Figure 1. It turns out that IR can improve the approximation ratio even further, from $3 - \frac{2}{n}$ to $2 - \frac{2}{n}$ [Meir et al. 2009; 2011]. Interestingly, we must know in advance whether our dataset is IR or not in order to apply the correct mechanism—selecting the mechanism after observing the data is no longer SP.

2.3 Generalizing from samples

A crucial requirement from supervised learning algorithms, and classification algorithms in particular, is that rules learned from sampled data can be applied to new data. Formally, we want the *empirical error* (on the dataset) to be close to the real error (measured on the entire distribution). Unfortunately, the SP requirement w.r.t. the real error cannot be obtained even if there is only one agent, due to the small chance that the sample does not reflect the agent's true opinion. In such cases we need to make some assumptions on the behavior of the agents. The *truthful approach* asserts that agents will only lie if they gain at least ϵ from doing so. In contrast, the *pure rationality approach* assumes that an agent will use a dominant strategy when one is available to him. Notably, for concept classes of a bounded VC dimension, all the algorithms mentioned above can be applied to sampled data under either of these assumptions, guaranteeing an approximation ratio that is arbitrarily close to $3 - \frac{2}{n}$.

3. A UNIFIED APPROACH TO SP CLASSIFICATION AND MECHANISM DESIGN

In addition to the technical advances mentioned above, many conceptual links have been drawn between the broad framework of mechanism design without money (i.e., without payments), and the problem of SP classification (the case of shared inputs in particular).

3.1 Judgment aggregation

A fairly intuitive connection is with the problem of *judgment aggregation* (JA) [Dokow and Holzman 2010]. In JA there is an agenda consisting of several logical expressions, and each agent has some opinion over the agenda. The different opinions, or judgments, should be aggregated to a single consistent assignment to all logical atoms. A simple mapping between the problems would identify each issue of the agenda with a data point, where every assignment vector corresponds to a binary classifier. The set of all legal (consistent) assignments then corresponds to the concept class of the learning problem. Note that in JA there is usually a requirement that the opinion of each agent itself be logically consistent. This requirement coincides with the IR requirement in the classification setting.

3.2 Facility location

In the *facility location* (FL) problem, agents report their location (usually in some metric space), and the mechanism outputs a location for a facility that is close, on average, to all agents [Procaccia and Tennenholtz 2009].

Consider a dataset labeled by several agents, and a binary cube where each dimension corresponds to a data point. We can now identify the label vector of each agent with a specific vertex of this cube. Similarly, any concept class (which defines the allowed labeling) corresponds to a set of vertices that constitutes the allowed locations. The IR condition in the FL setting is translated to the restriction that agents' locations are limited vertices where the facility can be placed. Moreover, the optimal location in FL corresponds to the ERM classifier.

4. CONCLUSION AND FUTURE DIRECTIONS

The above correspondences imply that questions of incentives and truthfulness in the distinct settings of JA, FL and learning can be studied in a unified model.

The design of learning algorithms that preclude or handle strategic behavior is advancing quickly, but certain obstacles still hinder its successful application to problems in the real world. First, the current models are quite general, overlook many intricacies that are featured in data from particular domains, and focus on worst-case analysis. Second, strong requirement of strategyproofness constrains the possible set of algorithms, whereas weaker strategic requirements may allow for much better results.

The first obstacle should be tackled with the help of experimental and empirical analysis of real data. As for the second, we believe that the emerging integration with the wider area of mechanism design will supply the necessary conceptual and technical tools to develop the proper solution concepts.

REFERENCES

Dokow, E. and Holzman, R. 2010. Aggregation of binary evaluations. *Journal of Economic Theory 145*, 495–511.

Meir, R., Almagor, S., Michaely, A., and Rosenschein, J. S. 2011. Tight bounds for strategyproof classification. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. Taipei, Taiwan, 319–326.

Meir, R., Procaccia, A. D., and Rosenschein, J. S. 2008. Strategyproof classification under constant hypotheses: A tale of two functions. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI)*. 126–131.

MEIR, R., PROCACCIA, A. D., AND ROSENSCHEIN, J. S. 2009. Strategyproof classification with shared inputs. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*. 220–225.

MEIR, R., PROCACCIA, A. D., AND ROSENSCHEIN, J. S. 2010. On the limits of dictatorial classification. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 609–616.

MEIR, R., PROCACCIA, A. D., AND ROSENSCHEIN, J. S. 2011. Algorithms for strategyproof classification. manuscript.

PROCACCIA, A. D. 2008. Towards a theory of incentives in machine learning. *ACM SIGecom Exchanges 7,* 2.

PROCACCIA, A. D. AND TENNENHOLTZ, M. 2009. Approximate mechanism design without money. In *Proceedings of the 10th ACM Conference on Electronic Commerce (ACM-EC)*. 177–186.

Prediction Markets for Education: An Experimental Study

CALI MORTENSON ELLIS
Departments of Public Policy and Political Science, University of Michigan
and
RAHUL SAMI
School of Information, University of Michigan

In this letter, we report the results of a quasi-experimental study of prediction markets as a pedagogical tool in an undergraduate setting.

Categories and Subject Descriptors: K.3.1 [**Computing Milieux**]: Computer Uses in Education—*Collaborative Learning*

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: prediction markets

In this letter, we summarize our recent work on using a popular forecasting tool, prediction markets, to supplement classroom learning. A full description of the experiment and results is forthcoming [Ellis and Sami 2012]; a short preliminary version appeared in the proceedings of the Computer Supported Collaborative Learning (CSCL) 2011 conference [Ellis and Sami 2011]. In contrast to most prior empirical research on prediction markets, which has focused on market outcomes and accuracy, we concentrate on the effect of the market on the traders themselves, as well as on characterizing the self-selected group of traders within the larger group of potential traders. Thus, these results may also be of interest to practitioners outside of the educational domain.

Prediction markets are widely used as forecasting tools, in a variety of commercial and non-commercial settings. In arguing for the use of prediction markets, proponents emphasize that they provide incentives that motivate traders to "ferret out accurate information" and "not amplify individual errors, but eliminate them" [Sunstein 2006]. These strengths match our goals as instructors: we want to train our students to search for relevant information, and critically analyze received information. Prediction markets also fit within the larger trend of integrating more interactive and technological resources into classroom learning. In order to test the performance of prediction markets as a learning tool, we carried out a controlled semester-long experiment in an introductory undergraduate political science class.

We used a nonequivalent comparison group quasi-experimental design using both control groups and pretests as per Shadish, Cook and Campbell [Shadish et al. 2002]. Half the class (traders) was randomly selected and granted permission to trade in the markets, while the other half served as the control. At the start and

Author's addresses: rsami@umich.edu, cmortens@umich.edu

the end of the course, we administered in-class surveys and quizzes. The surveys included questions to gauge student demographics, interest in the class topic, and reading behavior. In addition, we included questions from the Motivated Strategies for Learning Questionnaire (MSLQ) [Pintrich et al. 1993]; these questions have been developed in the education literature to assess students' skill at self-regulating their own learning. The quizzes contained general-knowledge questions related to the market topics. Our final complete dataset included data on 129 students.

Our market deployment built on pilot studies from two previous semesters to increase student comprehension, participation, and trading activity. We deployed a market that used the Zocalo open-source market software *(zocalo.sourceforge.net)*, with a custom user interface developed using the Drupal application platform *(www.drupal.org)*. Due to the thin market setting, we used the market scoring rule market maker introduced by Hanson [2003]. We created 12 markets relevant to the topic of the course, world politics. In order to avoid impacting grades (even indirectly) with our randomized experiment, we chose market topics to be tangential, although still relevant, to the syllabus of the class. At the end of the semester, students could cash in their trading budgets based on their performance for a modest amount of money (about $10).

Our first set of results are based on comparisons between the trading and control groups. Unfortunately, we did not find evidence that the trading group had significantly greater improvements in enthusiasm for the subject, independent reading, or quiz knowledge. In fact, all students reported lower levels of enthusiasm for the subject of the course at the end of the semester; one factor may be that the end of semester survey occurred shortly before the final exam. The entire class showed statistically significant improvement in their quiz score results. In both cases, however, the differences between traders and control was not significant.

Our second set of results makes a finer distinction between those who were randomized into the trader group but did not choose to trade (inactive traders) and those who chose to trade more actively in the markets (active traders). There were 22 inactive traders and 45 active traders. One striking finding was that that active traders had a higher level of broad information gathering: At the start of the semester, 82% of those students who would eventually become active traders reported reading about the politics of other countries at least once per week; this was significantly higher than inactive traders (54%) and the overall class (70%). Active traders also had higher average MSLQ scores, indicating that they were more self-motivated learners. We noted that active traders' quiz scores improved the most of any group, although the differences-in-differences with inactive traders were not statistically significant.

We detected a possible gender bias among active traders which may be worth considering when deploying prediction markets in a classroom setting. The class was 44 percent female, with no statistically significant difference between students randomized into control and treatment groups. Of the 45 active traders, 29 were male and 16 were female. The proportion of female active traders was slightly (but not significantly) lower than the trader group as a whole. In terms of number of trades, however, there is a statistically significant difference: The average number of trades for male traders was 15.5, but the average for female traders was 6.1.

Active traders indicated a high level of enthusiasm for the use of prediction markets in class. Of the active traders, 68 percent reported that their reason for participating was that they "wanted to win money," and 44 percent were "interested in learning about the topics."

We also gathered information on what they relied on for their trading decisions. 58% reported that trading decisions were made "based on personal beliefs," followed by 51% based on news reports. The smallest number of students reported making trades based on the outcome they wanted (4 percent) or based on the trades of others (i.e. the price reported on the graph - 6 percent). The results of our studies yield several new insights about the use of prediction markets as learning tools.

While we found no significant improvement in students' enthusiasm or extent of topical reading, we did find that those already reading broadly at the course start were more likely to trade actively in the markets, and those who did trade actively reported that they enjoyed the addition of markets to the class. These results taken together indicate that the prediction markets may be best deployed in a classroom of students who are highly motivated and already engaged in the subject matter. An elective upper-level undergraduate course or a graduate course may be more appropriate settings for using prediction markets as an educational tool. Further, instructors should be aware of the possibility of gender-biased participation One limitation of our study is that we intentionally picked market questions that were tangential to the course content. An important direction for future research is to study (in a non-randomized design) the use of prediction markets that are closer to the core of the syllabus as a tool for student engagement.

Acknowledgements:

This research was partially supported by the National Science Foundation under grant CCF-0728768. We are very grateful to Allan Stam for supporting these experiments in his courses; Skip Lupia, Stephanie Teasley, Kara Makara, and Rick Wilson for numerous helpful suggestions and feedback; Stan Dimitrov for assistance with the market interface; and Mohammed Hadhrawi, Yung-Ju Chang, and Yi-Chin Wu for helpful discussions on an earlier project.

REFERENCES

Ellis, C. and Sami, R. 2011. Learning with prediction markets: An experimental study. In *Proceedings of the 9th International Conference on Computer Supported Cooperative Learning (CSCL '11).*

Ellis, C. and Sami, R. 2012. Learning political science with prediction markets: An experimental study. *PS: Political Science and Politics (forthcoming).*

Hanson, R. 2003. Combinatorial information market design. *Information Systems Frontiers 5,* 1, 107119.

Pintrich, P., Smith, D., Garcia, T., and McKeachie, W. 1993. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement 53,* 3, 801–813.

Shadish, W. R., Cook, T. D., and Campbell, D. T. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Houghton-Mifflin, Boston.

Sunstein, C. R. 2006. Deliberating Groups versus Prediction Markets (or Hayek's Challenge to Habermas). *Episteme: A Journal of Social Epistemology 3.3*, 192–213.

Stability and Competitive Equilibrium in Matching Markets with Transfers

JOHN WILLIAM HATFIELD
Graduate School of Business, Stanford University
and
SCOTT DUKE KOMINERS
Becker Friedman Institute, University of Chicago

This note surveys recent work in generalized matching theory, focusing on trading networks with transferable utility. In trading networks with a finite set of contractual opportunities, the substitutability of agents' preferences is essential for the guaranteed existence of stable outcomes and the correspondence of stable outcomes with competitive equilibria. Closely analogous results hold when venture participation is continuously adjustable, but under a concavity condition on agents' preferences which allows for some types of complementarity.

Categories and Subject Descriptors: J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*; K.4.4 [**Computers and Society**]: Electronic Commerce

General Terms: Economics, Theory

Additional Key Words and Phrases: Matching, Networks, Joint Ventures, Stability, Competitive Equilibrium, Core, Efficiency

1. INTRODUCTION

In the half-century since Gale and Shapley [1962] introduced the stable marriage model, matching theory has been extended to encompass successively more general economic settings with relationship-specific utilities. The fundamental solution concept in this literature is *stability*, the condition that no group of agents can *block* the match outcome by recontracting. Stable outcomes have been shown to exist in two-sided matching markets—including those for which the matching process determines contractual terms in addition to partnerships—even when agents on both sides of the market may match to multiple agents on the other side. Crucial for these results, however, is a *substitutability* condition on agents' preferences, which requires that when an agent is presented with new matching opportunities, that agent never desires a previously-rejected opportunity.[1] The existence results

[1]Substituable preferences are sufficient and necessary for the existence of stable outcomes in settings of many-to-one matching (Roth [1984] proved the sufficiency result; Hatfield and Kojima [2008] proved the necessity result), and in settings of many-to-many matching with and without contracts (Roth [1984], Echenique and Oviedo [2006], Klaus and Walzl [2009], and Hatfield and Kominers [2011a] proved sufficiency results; Hatfield and Kojima [2008] and Hatfield and Komin-

Authors' addresses: `hatfield_john@gsb.stanford.edu`, `skominers@uchicago.edu`

This article describes the results of [Hatfield et al. 2011] and [Hatfield and Kominers 2011b].

The authors appreciate the helpful comments of Fuhito Kojima, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp.

(along with associated structural characterizations) for two-sided matching models extend to the more general setting of multi-stage, vertical "supply-chain" networks, so long as agents' preferences over the objects being traded are substitutable.[2]

Matching theory now has a number of high-profile applications[3]; hence, it is important to understand the possibilities and limitations of matching theory for market design. In our previous work [Hatfield and Kominers 2012], we showed that supply-chain market structure is essentially necessary for the guaranteed existence of stable outcomes in matching markets without transfers: If a market does not exhibit supply chain structure, then there exists some agent who may both buy from and sell to another agent (perhaps by way of intermediaries); if that agent has additional trading opportunities, then there exist substitutable preferences for all agents such that no stable outcome exists.

The recent work of [Hatfield et al. 2011] and [Hatfield and Kominers 2011b] shows that it is possible to accommodate more general network structures when utility is quasilinear in a transferable numeraire. In such settings, the underlying nature of contractual relationships is important: Substitutability remains essential in trading networks with a finite set of contractual opportunities, while some complementarities may be allowed when venture participation is continuously adjustable.

2. TRADING NETWORKS WITH BILATERAL CONTRACTS

The work of [Hatfield et al. 2011] considers contracting over *trades*—each of which specifies a buyer, a seller, and terms of exchange—and augments the contractual set with transfers of a continuously divisible numeraire over which agents' utilities are assumed to be quasilinear.[4] In this setting, the choice-theoretic notion of substitutability described above is equivalent to the demand-theoretic substitutability condition that when the prices an agent faces rise, that agent chooses any previously-chosen purchase opportunities for which prices are unchanged and

ers [2011a] proved necessity results). Meanwhile, Hatfield and Milgrom [2005] showed that substitutable preferences are sufficient for the existence of stable outcomes in the setting of many-to-one matching with contracts, but Hatfield and Kojima [2008] showed that there is no corresponding necessity result (see also [Hatfield and Kojima 2010]).

[2]Ostrovsky [2008] and Hatfield and Kominers [2012] showed that substitutable preferences are sufficient to guarantee the existence of stable outcomes; Hatfield and Kominers [2012] proved a corresponding necessity result.

[3]The theory of matching without contracts has been used to design the National Resident Matching Program [Roth and Peranson 1999], the gastroenterology match [Niederle and Roth 2003; 2005; McKinney et al. 2005], and school choice programs in New York [Abdulkadiroğlu et al. 2005; 2009] and Boston [Abdulkadiroğlu et al. 2005]. The theory of matching with contracts has been used to analyze the impact of "branch-of-choice" contracts on cadet–branch matching [Sönmez and Switzer 2011; Sönmez 2011]. In addition, matching with contracts has recently been used as a technical tool for understanding markets with budget-constrained buyers [Hatfield and Milgrom 2005; Ashlagi et al. 2010] (see also [Alaei et al. 2011a; 2011b]), matching markets with regional caps [Kamada and Kojima 2011; 2012], markets with differentiated goods and price controls [Hatfield, Plott, and Tanaka 2011; 2012], and matching with minimal quotas [Fragiadakis et al. 2011].

[4]This model generalizes the models of Crawford and Knoer [1981], Kelso and Crawford [1982], Gul and Stacchetti [1999; 2000], and Sun and Yang [2006; 2009].

rejects any previously-rejected sale opportunities for which prices are unchanged.[5]

In this setting, a stable outcome is a set of contracts that is both individually rational (for all agents) and unblocked. In the presence of substitutable preferences, an extension of the [Kelso and Crawford 1982] *salary-adjustment process* shows that stable outcomes exist in trading networks. Such outcomes are in the core (and, hence, are efficient)[6]; moreover, the set of stable outcomes is essentially equivalent to the set of competitive equilibria. The space of substitutable preferences is the maximal domain over which the existence of stable outcomes may be guaranteed—that is, for any domain of preferences strictly larger than that of substitutability, the existence of competitive equilibria and stable outcomes cannot be guaranteed.[7]

3. MULTILATERAL CONTRACTING

Despite the maximal domain results described in Section 2, there are a number of economic settings for which preference substitutabilty is not a valid assumption: automobile manufacturing requires compelementary inputs for production [Fox 2008]; advertising campaigns are coordinated across multiple publishers; information technology firms collaborate on multiparty joint research ventures. In [Hatfield and Kominers 2011b], we introduce a *multilateral matching* framework which allows us to analyze the aforementioned economic environments. In multilateral matching, sets of two or more agents may enter into contracts over participation in multilateral *ventures* such as coordinated production or joint research. Certain forms of complementarity can be expressed through multilateral contracts; in particular, the multilateral matching framework embeds a large class of economies with production complementarities.

The maximal domain result of [Hatfield et al. 2011] implies that the existence of stable multilateral contracting outcomes cannot be guaranteed when venture participation is discrete. Nevertheless, stable multilateral contracting outcomes do exist when venture participation is continuously adjustable, so long as agents' preferences are concave in venture participation and there exists a numeraire over which agents' preferences are quasilinear.[8] Furthermore, stable outcomes correspond to competitive equilibria when agents' utilities are concave. Conversely, and in close analogy with the results of [Hatfield et al. 2011], competitive equilibria induce outcomes

[5]Theorem 1 of [Hatfield et al. 2011] shows that this definition of substitutability is equivalent to submodularity of the indirect utility function; this generalizes results of Gul and Stacchetti [1999] and Sun and Yang [2009]. This result also corresponds to an analogous result of Hatfield and Kominers [2012], which shows that in settings without transfers, substitutability is equivalent to *quasi*submodularity. Substitutability can also be characterized in terms of $M^\natural$-concavity of the utility function; see [Reijnierse et al. 2002; Fujishige and Yang 2003].

[6]By contrast, in settings without transfers, stable outcomes need not be in the core; see [Blair 1988; Echenique and Oviedo 2006]. In the setting of Ostrovsky [2008], Westkamp [2010] characterized the class of network structures for which stable outcomes are guaranteed to be efficient and in the core.

[7]Formally, this means that when one agent's preferences are not substitutable, there exist substitutable preferences for all other agents such that neither stable outcomes nor competitive equilibria exist. (In fact, the substitutable preferences for the other agents can be chosen to have a particularly simple form; see Theorem 8 of [Hatfield et al. 2011].)

[8]The concavity assumption is natural in markets with decreasing returns to scale and scope, but is violated in settings with fixed costs.

that are stable and in the core. Analogues of the first and second welfare theorems hold as well, showing in particular that stable outcomes and competitive equilibria are efficient. We extend the model to allow for cross-contract externalities; even when such externalities are introduced, competitive equilibria exist (although they may not be efficient).

4. DISCUSSION

A crucial distinction between the models described in Sections 2 and 3 lies in the conditions on preferences necessary for the main results: substitutability is essential in markets with discrete trades, whereas concavity is essential in markets with continuously adjustable participation levels. Under their respective key conditions, however, these models are surprisingly parallel:[9] stable outcomes exist, are in the core (and hence are efficient), and correspond to competitive equilibria. We hope that future work will yield insight into these parallels through a deeper understanding of the relationship between substitutability and concavity.

The importance of preference conditions for the guaranteed existence of stable outcomes suggests that market design may be difficult in settings where these conditions are violated. Recent large market results of Kojima et al. [2010], Ashlagi et al. [2011], and Azevedo et al. [2011] suggest that these concerns could be mitigated by sufficient market thickness. For small markets, by contrast, new market design approaches may be needed.

REFERENCES

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. 2005. The New York City high school match. *American Economic Review 95*, 364–367.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. 2009. Strategyproofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. *American Economic Review 99*, 1954–1978.

Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. 2005. The Boston public school match. *American Economic Review 95*, 368–371.

Alaei, S., Jain, K., and Malekian, A. 2011a. Competitive equilibrium in two sided matching markets with general utility functions. Preprint, `arXiv:1006.4696v3`.

Alaei, S., Jain, K., and Malekian, A. 2011b. Competitive equilibrium in two sided matching markets with general utility functions. *SIGecom Exchanges 10,* 2, 34–36.

Ashlagi, I., Braverman, M., and Hassidim, A. 2011. Stability in large matching markets with complementarities. Mimeo, Massachusetts Institute of Technology. (2011 extended abstract, entitled "Matching with couples revisited". In *Proceedings of the 12th ACM Conference on Electronic Commerce*. 335–336.)

Ashlagi, I., Braverman, M., Hassidim, A., Lavi, R., and Tennenholtz, M. 2010. Position auctions with budgets: Existence and uniqueness. *B.E. Journal of Theoretical Economics – Advances 10*, Article 20.

Ausubel, L. M. 2006. An efficient dynamic auction for heterogeneous commodities. *American Economic Review 96*, 602–629.

Azevedo, E., Weyl, E. G., and White, A. 2011. Walrasian equilibrium without gross substitutes. Mimeo, University of Chicago.

Blair, C. 1988. The lattice structure of the set of stable matchings with multiple partners. *Mathematics of Operations Research 13*, 619–628.

[9]Similar parallels have been observed in the auction theory literature; see [Ausubel 2006].

Crawford, V. P. and Knoer, E. M. 1981. Job matching with heterogeneous firms and workers. *Econometrica 49*, 437–450.

Echenique, F. and Oviedo, J. 2006. A theory of stability in many-to-many matching markets. *Theoretical Economics 1*, 233–273.

Fox, J. T. 2008. Estimating matching games with transfers. *NBER Working Paper 14382*.

Fragiadakis, D., Iwasaki, A., Troyan, P., Ueda, S., and Yokoo, M. 2011. Strategy-proof mechanisms for two-sided matching with minimum and maximum quotas. Mimeo, Stanford University.

Fujishige, S. and Yang, Z. 2003. A note on Kelso and Crawford's gross substitutes condition. *Mathematics of Operations Research 28*, 463–469.

Gale, D. and Shapley, L. S. 1962. College admissions and the stability of marriage. *American Mathematical Monthly 69*, 9–15.

Gul, F. and Stacchetti, E. 1999. Walrasian equilibrium with gross substitutes. *Journal of Economic Theory 87*, 95–124.

Gul, F. and Stacchetti, E. 2000. The English auction with differentiated commodities. *Journal of Economic Theory 92*, 66–95.

Hatfield, J. W. and Kojima, F. 2008. Matching with contracts: Comment. *American Economic Review 98*, 1189–1194.

Hatfield, J. W. and Kojima, F. 2010. Substitutes and stability for matching with contracts. *Journal of Economic Theory 145*, 1704–1723.

Hatfield, J. W. and Kominers, S. D. 2011a. Contract design and stability in matching markets. Mimeo, Harvard Business School.

Hatfield, J. W. and Kominers, S. D. 2011b. Multilateral matching. Mimeo, University of Chicago. (2011 extended abstract. In *Proceedings of the 12th ACM Conference on Electronic Commerce*. 337–338.)

Hatfield, J. W. and Kominers, S. D. 2012. Matching in networks with bilateral contracts. *American Economic Journal: Microeconomics*. (2010 extended abstract. In *Proceedings of the 11th ACM Conference on Electronic Commerce*. 119–120.)

Hatfield, J. W., Kominers, S. D., Nichifor, A., Ostrovsky, M., and Westkamp, A. 2011. Stability and competitive equilibrium in trading networks. Mimeo, Stanford University.

Hatfield, J. W. and Milgrom, P. 2005. Matching with contracts. *American Economic Review 95*, 913–935.

Hatfield, J. W., Plott, C. R., and Tanaka, T. 2011. Price controls, non-price quality competition, and nonexistence of competitive equilibrium. Mimeo, Stanford University.

Hatfield, J. W., Plott, C. R., and Tanaka, T. 2012. Understanding price controls and non-price competition through matching theory. *American Economic Review Papers & Proceedings*.

Kamada, Y. and Kojima, F. 2011. Improving efficiency in matching markets with regional caps: The case of the Japan Residency Matching Program. Mimeo, Stanford University.

Kamada, Y. and Kojima, F. 2012. Stability and strategy-proofness for matching with constraints: A problem in the Japanese medical matching and its solution. *American Economic Review Papers & Proceedings*.

Kelso, A. S. and Crawford, V. P. 1982. Job matching, coalition formation, and gross substitutes. *Econometrica 50*, 1483–1504.

Klaus, B. and Walzl, M. 2009. Stable many-to-many matchings with contracts. *Journal of Mathematical Economics 45*, 422–434.

Kojima, F., Pathak, P. A., and Roth, A. E. 2010. Matching with couples: Stability and incentives in large markets. Mimeo, Harvard Business School.

McKinney, C. N., Niederle, M., and Roth, A. E. 2005. The collapse of a medical labor clearinghouse (and why such failures are rare). *American Economic Review 95*, 878–889.

Niederle, M. and Roth, A. E. 2003. Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match. *Journal of Political Economy 111*, 1342–1352.

Niederle, M. and Roth, A. E. 2005. The gastroenterology fellowship market: Should there be a match? *American Economic Review 95*, 372–375.

Ostrovsky, M. 2008. Stability in supply chain networks. *American Economic Review 98*, 897–923.

Reijnierse, H., van Gellekom, A., and Potters, J. A. M. 2002. Verifying gross substitutability. *Economic Theory 20*, 767–776.

Roth, A. E. 1984. Stability and polarization of interests in job matching. *Econometrica 52*, 47–57.

Roth, A. E. and Peranson, E. 1999. The effects of the change in the NRMP matching algorithm. *American Economic Review 89*, 748–780.

Sönmez, T. 2011. Bidding for army career specialties: Improving the ROTC branching mechanism. Mimeo, Boston College.

Sönmez, T. and Switzer, T. B. 2011. Matching with (branch-of-choice) contracts at United States Military Academy. Mimeo, Boston College.

Sun, N. and Yang, Z. 2006. Equilibria and indivisibilities: gross substitutes and complements. *Econometrica 74*, 1385–1402.

Sun, N. and Yang, Z. 2009. A double-track adjustment process for discrete markets with substitutes and complements. *Econometrica 77*, 933–952.

Westkamp, A. 2010. Market structure and matching with contracts. *Journal of Economic Theory 145*, 1724–1738.

Contingency Exigency

DANIEL REEVES
Beeminder

Lawyers and debt collectors are classic examples of people you hire on contingency. The idea is to hire someone with payment contingent on the fruits of their labor. You don't pay them if you don't win, i.e., if you yourself don't get paid. And if you luck out, they share in the windfall.

The agency problem here is insurmountable. If you pay them hourly then they have incentive to drag things out. If you pay them a percentage of the winnings then, unless the percentage is unreasonably high, you have the opposite problem: they'll have incentive to skimp on the time they put in.

So we'll ignore the incentives and assume that this agent[1] will conscientiously put in exactly as much time as they would if they were you.

The question: What's a fair payment function? We'll take fairness to mean that your agent earns in expectation exactly what they would've earned had you paid a straight hourly rate for their time. Assume that you and your agent agree on what that non-contingency rate, r, would be, and also that you agree on the probability distribution, F, of the payout (e.g., the lawsuit settlement, or the debt collected). So we seek a function, $\omega(t, X)$ that takes the number of hours the agent spent, t, and the eventual realized payout, X, and returns the amount you should pay the agent, parametrized by r and F.

The fairness desideratum can be expressed as $E_F[\omega(t, X)] = r \cdot t$. We also impose two other reasonable constraints: $\omega(t, 0) = 0$ for all t. No payout means no wages. And $\omega(t, X)$ is linear in t. Putting in twice the hours means twice the wages. Finally, assume $r \cdot t < E_F[X]$, namely, the agent doesn't do more work than the payout is, in expectation, worth.

Find ω. How would you relax the first constraint, to ensure a bare mininum payment, m, or $m \cdot t$, regardless of payout?

Warning: Do not overthink this. This puzzle sacrifices difficulty for practicality. If you wanted to work out a fair payment arrangement with a trusted friend as your agent — and the uncertainty of both the time requirement and the payout precluded the simplest options, like a flat percentage or a straight hourly payment — what would you actually do?

Send solutions to the puzzle editor at `dreeves@beeminder.com` with subject: `conex`. The author(s) of the most elegant solution (as judged by the editor) will be allowed to publish it in the next issue of the Exchanges (ties broken in favor of earlier submissions). To make the solutions accessible to a wide audience, please try to

[1] I'm using the term "agent" in the legal (and colloquial) sense of someone who acts on another's behalf.

minimize technical jargon. Until the winner is chosen the editor will not give any hints or feedback.

Solution to Exchanges 10.1 Puzzle: Baffling Raffling Debaffled

PRESTON MCAFEE
Yahoo! Research

[Puzzle Editor's Note: This is the winning solution to *Baffling Raffling* from issue 10.1. The mechanism described there is sometimes known as a Chinese Auction. It is also equivalent, as McAfee points out, to a special case of a Cournot problem. An alternative formulation is: I decide a bid x, pay it in full, and then win the good with probability x/X where X is the sum of all the bids. Generalizing the question in the original puzzle, this solves the game for an arbitrary vector of common-knowledge valuations, i.e., the complete-information case with n agents.]

Notation: x_i is i's bid (the number of tickets bought by i) and v_i is the value of i, indexed so that $v_1 \geq v_2 \geq \ldots$. Let

$$X_{-i} = \sum_{j \neq i} x_j \text{ and } X = \sum_j x_j.$$

First, note that the payoff to i, given the choices of others, is $\dfrac{x_i}{x_i + X_{-i}} v_i - x_i$. The choice of x_i is restricted to $x_i \geq 0$, and probably should be restricted to integers. I will ignore this constraint. [This turns out to be moot for the specific (carefully constructed) valuations given in the puzzle.] Note that the individual maximization problem is equivalent to maximizing

$$\frac{x_i}{x_i + X_{-i}} - \frac{1}{v_i} x_i \equiv p(X) x_i - c_i x_i,$$

where $p(X) = \frac{1}{X}$ and $c_i = \frac{1}{v_i}$. The solution to the problem is just the solution to the standard constant marginal cost Cournot problem, with a unitary elasticity demand curve and asymmetric firms. While this is a common graduate student exercise, the solution isn't necessarily well-behaved.

To characterize the equilibria, return to the profit functions $\frac{x_i}{x_i + X_{-i}} - c_i x_i$. This function is concave, so the first order conditions characterize the maximum. The first derivative is $\frac{X_{-i}}{(x_i + X_{-i})^2} - c_i = \frac{X - x_i}{X^2} - c_i$. As the values of c_i increase in i (being the reciprocals of the v's), there will be a value n so that the first n have $x_i > 0$ and all others have $x_i = 0$. Note that all the agents with positive production have a zero first order condition, or $\frac{X - x_i}{X^2} - c_i = 0$. Summing these gives

$$0 = \frac{nX - X}{X^2} - \sum_{i=1}^{n} c_i,$$

which solves for

$$\frac{1}{X} = \frac{1}{n-1} \sum_{i=1}^{n} c_i,$$

and note immediately from the first order conditions that $n > 1$. An equilibrium has been achieved if, given this value of X, the first n firms want to enter and produce positive amounts and no others do, which is equivalent to

$$\frac{1}{X} - c_n \geq 0 \geq \frac{1}{X} - c_{n+1} \text{ or}$$
$$c_n \leq \frac{1}{X} \leq c_{n+1} \text{ or}$$
$$c_n \leq \frac{1}{n-1}\sum_{i=1}^{n} c_i \leq c_{n+1}.$$

Once we have an equilibrium number of agents and $\frac{1}{X} = \frac{1}{n-1}\sum_{i=1}^{n} c_i$, we can use the first order conditions $0 = \frac{X-x_i}{X^2}$, or $x_i = X - c_i X^2$ to generate the number of tickets purchased.

Using the Mathematica functions below, that yields $\langle 119, 77, 21, 0\rangle$ with profits of $\langle 144.5, 42.35, 2.25, 0\rangle$.

Is the solution unique? Let $p_n = \sum_{i=1}^{n} c_i$. The coumputation given shows

$$c_n \leq p_n$$
$$\iff c_n \leq p_{n-1}$$
$$\implies c_{n-1} \leq p_{n-1}.$$

Thus take the largest equilibrium n^{**}. For all k smaller,

$$c_k \leq p_k.$$

But consider any hypothetical smaller equilibrium n^*. As shown it satisfies

$$p_{n^*+1} \leq c_{n^*+1}$$

This would be a contradiction except for ties. If the c's were strictly increasing we would have the first inequality strictly and be done. If some c's are equal, the additional firms/agents produce/bid zero (since we are satisfying the price inequality with equality) and can be safely ignored.

The final question: how did the profits compare? The profit vector (seller, buyers) was $\langle 336.35, 144.15, 0, 0, 0, 0\rangle$, and under the raffle it is $\langle 217, 144.5, 42.35, 2.25, 0, 0\rangle$. So Nora gained the most.

Implementation of the solution in Mathematica

Following is Mathematica code to compute the equilibrium bids and profits. The helper function `bz` gives the hypothetical equilibrium bids (as a function of the vector of values) if we knew all agents would, in equilibrium, participate. Another helper function, `bs`, gives the equilibrium bids, without assuming full participation, if the values are in ascending order, which is of course WLOG. The `bs` function works by recursively re-solving for equilibrium bids with the subset of agents for which `bz` yields positive bids. Finally, `bids` gives the equilibrium bids for arbitrary values (by just sorting, calling `bs`, and then unsorting). Additionally, `prof` gives the expected profit to each agent in equilibrium.

```
bz[v_]:= With[{n = Length[v], r = Total[1/v]}, (n-1)(r-(n-1)/v)/r^2]
bs[v_]:= With[{x = bz[v]}, If[x[[1]]<0, Prepend[bs[Rest[v]], 0], x]]
bids[v_]:= bs[Sort@v][[Ordering@Ordering@v]]
prof[v_]:= With[{b = bids[v]}, v*b/Total[b] - b]
```